

The increasing opacity of gene therapy

The influence of commercial interests is highlighted by recent reports of deaths during trials of gene therapies. The loosening of public scrutiny of such studies needs to be reconsidered.

A decade ago, when gene therapy burst on the scene promising potential cures for all manner of diseases, researchers made a shrewd political choice. By submitting to the notion that studies of a highly experimental therapy should be publicly scrutinized, they bought the goodwill of the American people and their trust that, whatever the bumps in the road, scientists and those who supported them were committed to patients' safety.

Ominously for gene therapy's public image, a more secretive tendency seems to be taking root among some gene therapists and their sponsors. It has emerged that at least six recent deaths in patients participating in gene-therapy trials have not been reported to the National Institutes of Health (NIH), despite federal guidelines requiring such reporting. Separately, the drug company Schering-Plough reported numerous serious side effects in cancer patients, but insisted that they be kept confidential (see *Nature* **402**, 6; 1999).

The deaths do not include the only one so far definitively linked to gene therapy, that of 18-year-old Jesse Gelsinger, who died in September at the University of Pennsylvania. Commendably, researchers there disclosed details fully and promptly (see *Nature* **401**, 517; 1999).

The six unreported deaths, documented on 3 November by *The Washington Post*, occurred in trials being run by Ronald Crystal and Jeffrey Isner, of Cornell Medical Center in Manhattan and Tufts University in Boston, respectively. The founders of separate biotechnology companies, the two are competing fiercely in a race to use gene therapy to grow new cardiac blood vessels. Both defend their decisions not to notify the NIH of the deaths, arguing that they were attributable not to gene therapy, but to the patients' underlying illnesses. Traditionally, that is a determination that researchers have left to the NIH's Recombinant DNA Advisory Committee (RAC), whose mandate is to

debate the ethics of gene therapy in public. Sponsors and investigators are also required to report deaths and adverse events to the Food and Drug Administration (FDA), and those involved in these cases say they did so. But the FDA keeps such reports confidential.

Why should researchers and companies be challenging established practice in a field that up to now has prided itself on its transparency? The answers are unsettling.

First, the field, which initially was three-quarters NIH-funded, is now heavily underwritten by industry, and the corporate concern for commercial secrecy has begun to make itself felt. Companies loathe disclosing anything that could hurt share prices. But corporate strong-arming of the RAC goes further: at least one company has pressured NIH officials to keep entire protocols off the public record. Fortunately, it has so far failed. But corporations may be finding a way around their problem. More and more companies — albeit still only a fraction of those involved — are funding their experiments entirely privately, which means they can evade RAC reporting requirements completely.

There is another troubling aspect to this new secrecy. Three years ago, Harold Varmus, the NIH director, stripped the RAC of its power to approve gene-therapy protocols. That left the RAC as an important, but less powerful, venue for public debate on the experiments (see *Nature* **384**, 297; 1996). The real power to determine their fate was left entirely behind closed doors at the FDA.

It is difficult not to conclude that the RAC's new toothlessness has left researchers and companies feeling too relaxed about their ability to flout its guidelines with impunity. If this is the outcome of what was essentially a compromise by Varmus — industry had been agitating for the RAC's complete abolition, and activists for its retention — perhaps it is time that that decision was revisited. ■

A cold wind from the East

Charges faced by a Ukrainian scientist could have a chilling effect on the future of international collaboration.

Life has been difficult for scientists in states of the former Soviet Union since the fall of communism. Those who have not left to work elsewhere have often only been able to maintain active research through collaboration with Western researchers. But even this, it seems, carries its risks. As a result of his participation in various European and US programmes, Sergei Piontkovski, a Ukrainian oceanographer who works at the Institute of Biology of the Southern Seas in Sebastopol, is facing charges of communicating 'secret' information to the West and illegally handling foreign currency (see *Nature* **401**, 835 & **402**, 6; 1999, and page 110 in this issue).

There is always the possibility that, unknown to his foreign collaborators, Piontkovski has been involved in such nefarious activities. But to most of those familiar with him and with the case, his 'crime' seems to have been no more than exchanging research data with foreign research teams on topics such as the distribution of plankton, and receiving support in hard currency from Western funding agencies.

By one account, the actions against him are the result of purely local initiatives, not of a clamp-down by central government. Piontkovski's international activities have been formally approved by the Ukrainian authorities in Kiev. It is to the credit of the Ukrainian Academy of Sciences that it has spoken out in his defence, as well as that of colleagues who are also facing scrutiny by the local secret police.

But the affair should not be dismissed as a little local difficulty, as it has profound international implications. If the charges against Piontkovski are upheld by the court later this month, the result will inevitably have a chilling effect on scientific collaboration, not only with the Ukraine, but also with its neighbouring states, many of whose economies — and scientific communities — are already in a fragile enough state. Hopefully, reason will prevail, and any charges against Piontkovski and his colleagues will be dropped with the same speed with which they were initially made. ■





UK synchrotron
Hopes raised at
Daresbury on
siting decision
p111



No, minister
Hostile reaction to
Australian plan for
universities
p113



At risk
Researchers in peril
in the former
Soviet Union
p114



Online battle
Commercial
publishers unite to
compete
p115

US universities find that demand for botanists exceeds supply

San Diego

Plant systematists, once desperate for academic jobs, have become hot property, aggressively pursued by US universities seeking to be at the forefront of biodiversity research in the new millennium.

Universities have engaged in bidding wars for leading plant systematists as if they were prized professional athletes. Some academic leaders are now concerned about balancing universities' needs for plant systematists with the available supply. "There are far too few systematists to fill available faculty positions nationally," says Brent Mishler, director of the herbaria at the University of California at Berkeley.

Plant systematists conduct a range of research, from biodiversity to genetically engineered crops. They may study the impact of climate change on flora, or identify a tropical plant from which a gene can be removed to secure a desired characteristic for an agricultural crop.

The recent completion of a 'tree of life' — describing the evolutionary relationships of the Earth's green plants — by the Green Plant Phylogeny Research Coordination Group has also spurred interest in systematics, says bryologist Mishler, who is a principal investigator for the five-year project (see *Nature* **400**, 602; 1999).

Noting that the twenty-first century is expected to be the era of the environment, Kris Krishtalka, director of the University of Kansas Natural History Museum, says: "There is going to be an enormous demand for systematists. Universities that cut faculty and students in organismal biology in the 1970s and 1980s are going to be scrambling now."

Indeed, they already are. The website for the American Society of Plant Taxonomists, whose membership of 1,300 scientists includes many of the nation's systematists (a term now virtually synonymous with taxonomists), list 15 systematics jobs at universities, herbaria and museums — a radical change from a decade ago, when plant systematists were crying out for jobs.

"The excitement of molecular biology led to a steady reduction in the number of facul-



Back in favour: growth in environmental studies has rekindled demand for plant systematists.

ty whose primary interest was in organismal biology and ecology," says Alison Richard, a primatologist who is Yale University's provost and a vociferous advocate of systematics. "That is turning around. There is growing student interest. Environmental issues have gone from having faddish, tree-hugger status to centre stage, viewed now as serious and valued."

As part of this reassessment, Yale has recruited plant systematist Michael Donoghue to assume the G. Evelyn Hutchinson Professorship of Ecology and Evolutionary Biology on 1 July 2000. Donoghue will vacate the directorship of Harvard's herbaria —

with five million specimens — for Yale's new Environmental Sciences Facility, being built next to the Peabody Museum of Natural History. The facility is to wed a variety of environmental studies with state-of-the-art molecular labs and the Peabody's 400,000-specimen herbarium collection.

The wooing of Donoghue — a riveting lecturer with an encyclopaedic knowledge of the literature — came after a long courtship, in which Harvard counter-offered in an attempt to retain the popular botanist's services. "Michael has a tremendous opportunity at Yale," says Dan Hartl, chairman of Harvard's Department of Organismic and Evolutionary Biology. "He is a big loss" for Harvard, says Hartl, who noted that systematics now "is an extremely volatile area".

With the departure of Donoghue, other botanists moving elsewhere and retirements, the Harvard department is expected soon to be down by a half a dozen positions. This has concerned some doctoral graduates of the Harvard programme, such as Mishler, who says there "is a lack of appreciation of systematic biology" in some segments of the university campus. But Hartl says the "multiple recruitments" now under way provide

Australia 'tops per capita emissions'

Sydney

Australia has the highest level of per capita emissions of greenhouse gases. That was the finding of a report presented to the Australian Senate to coincide with last week's international meeting of environment ministers in Bonn.

Environmental economists at the Australia Institute used official returns to the United Nations to calculate 'carbon dioxide equivalents' by adding non-energy emissions from changes in land use, forestry and agriculture to emissions of the greenhouse gases — carbon dioxide, methane and nitrous oxide.

The research group ranked per capita emissions from the 35 developed nations that accepted targets to reduce emissions

under the Kyoto Protocol — known as Annex B countries.

Australia scored 26.7 tonnes of emissions per capita annually — double the average of the wealthy countries in Annex B — followed by the United States, Canada, New Zealand, Germany, the United Kingdom and Japan (9.5 tonnes).

Institute director Clive Hamilton says Australia has "exceptionally high emissions" from land use. Only Australia and the United Kingdom show land use as a net source of emissions. Because perceptions have been based on energy consumption, the report says, the analysis will affect the application of the 'polluter pays' principle, and exonerate Americans as the highest emitters per head of population. Peter Pockley

► “a great opportunity” to expand plant biology.

At the University of Arizona, the director of their herbarium, Lucinda McDade, is leaving to become herbarium curator at the Academy of Natural Sciences in Philadelphia.



Donoghue: offered jobs at both Harvard and Yale.

Eugene Sander, a biochemist who is dean of Arizona's College of Agriculture, says McDade had secured two National Science Foundation grants to modernize Arizona's collection of desert flora from the southwest

United States and northwest Mexico. “Our university is losing a fine scientist who will be difficult to replace,” says Sander.

The University of Alaska is having trouble recruiting a director for its herbarium of high-latitude specimens, which is important for research on global warming. According to David McGuire, chairman of the Alaska search committee, when the university advertised last July for an assistant professor, it received only 15 applications, half of what was desired; and that number was achieved after pleas to herbarium directors around the nation.

Alaska is interviewing three candidates, says evolutionary biologist McGuire. Beginning today, 11 November, the American Institute of Biological Sciences, an umbrella organization for 55 biological societies, will hold a summit in Virginia for organization presidents. It will address a range of planning issues, including the availability of plant systematists, says the institute's president, Gregory Anderson, a University of Connecticut botanist.

Approximately 150 university departments of ecology and evolutionary biology include plant systematists, Anderson says. About 30 to 50 students may graduate annually, with half seeking work at universities. Consideration should be given to increasing the number of trained plant systematists — particularly those with molecular biology skills — to balance supply and demand, he says.

It is important to address this balance soon, because many plant systematists are on the verge of retirement — for example Shirley Graham, a professor at Kent State University in Ohio and president of the American Society of Plant Taxonomists.

“My husband and I are the only botanists in our department of 25 faculty, and we are both a year from retirement,” says Graham. “Botanists have retired and been replaced by other disciplines. Now we have to rethink what constitutes a good programme in botany.”

Rex Dalton

Publishing group offers peer review on PubMed Central

Paris

The Current Science Group is to establish a new publishing house dedicated to providing peer review and other services for PubMed Central, the free repository for life-science papers that is due to be launched early in the new year by the US National Institutes of Health.

The new venture, called BioMed Central, will be established as a separate company within the group, with an initial staff of around 20. Other parts of the Current Science Group will contribute to the venture, says Vitek Tracz, the group's chief executive officer. “We have a tradition of working together as smallish, independent units with little bureaucracy, and we like it that way.”

The new company, whose ‘under construction’ website can be viewed at <http://www.biomedcentral.com/>, is intended to provide a “large multidisciplinary peer-review mechanism”, sitting on top of PubMed Central that “would be equivalent to thousands of journals”, says Tracz. “It is the most exciting and the most important publishing initiative I have got myself into.”

On acceptance, papers submitted to Bio-

Med Central would be deposited in PubMed Central, where they would be available free to all. Tracz believes that it will be feasible to give this primary literature away, and make money from a mix of advertising, sponsorship, e-commerce and subscription to editorial material such as news and reviews.

A similar logic is being followed by a venture launched last month by the Community of Science, a private US company that gives researchers information about funding opportunities and other activities (see *Nature* 401, 516; 1999).

Many publishers remain sceptical whether such Internet business models — largely untested in science publishing — will provide enough revenue to support quality publishing operations (see page 115).

This is partly a defensive reflex — some publishers have been making what many consider to be excessive profits (see *Nature* 397, 195–200; 1999). But there are genuine concerns among publishers and scientific societies that, if change is too rapid and ill-considered, the baby of quality literature might be thrown out with the bathwater.

Declan Butler

Ukraine denies arresting biologist

London

The Ukrainian government last week denied accusations that a prominent marine biologist, Sergei Piontkovski, had been arrested and charged with sending secret information abroad and handling foreign currency illegally.

The statement was made by the first deputy foreign minister of Ukraine in response to concerns expressed by the European Commission over the arrest and detention of research workers from the Institute of Biology of the Southern Seas (IBSS) in Sebastopol (see *Nature* 401, 835 & 402, 6; 1999).

Piontkovski and colleagues are back at work, but reports suggest that they still face charges, and Piontkovski's trial is scheduled for later this month. The security service continues to hold IBSS computers and documents. A spokesperson for the European aid programme INTAS said the Ukrainian Ministry of Science had been supportive in trying to persuade authorities of the innocence of IBSS staff. “It is clear to the authorities that Piontkovski was arrested on non-legal grounds,” said a spokesperson.

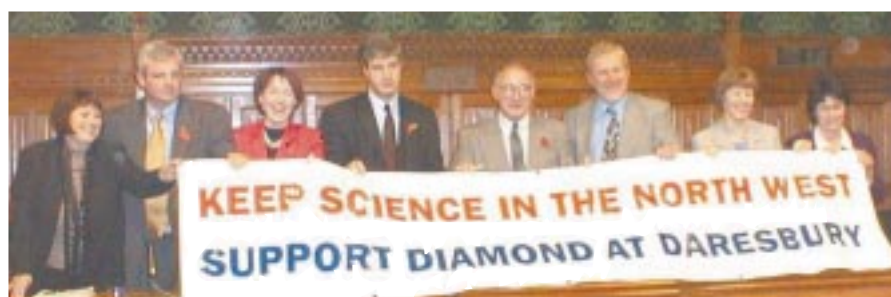
Grants from INTAS and the British government — on biodiversity — have been

the focus of inquiries by the security services. Last week, Yuri Tokarev, deputy director of the IBSS, who also faces criminal charges, said charges related to bioluminescence studies. This could be connected with the awarding of a grant by the US Office of Naval Research (ONR) to the Plymouth Marine Laboratory in the United Kingdom in August, for work with which Piontkovski is involved to create a database on bioluminescence in the oceans.

A US Navy spokesperson said: “This project is part of an international effort with broad participation and is expected to increase knowledge of the basic scientific properties of the oceans, including the distribution and abundance of bioluminescent animals and plants.”

He said it would be “inappropriate” for the navy to comment on the alleged arrest, and that ONR was a “leading supporter of oceanographic research, and sponsors many unclassified basic research and data-gathering projects around the world”. The head of IBSS's laboratory of biophysical ecology last week said his was the only laboratory in the former Soviet Union dealing with bioluminescence in the sea.

Natasha Loder



Out in force: local Members of Parliament show their support for siting Diamond at Daresbury.

Daresbury senses victory in battle for UK synchrotron

London

Speculation is mounting that Britain's next-generation synchrotron source, Diamond, will be located at Daresbury Laboratory outside Manchester. Scientists campaigning for Daresbury, rather than the rival Rutherford Appleton Laboratory (RAL) near Oxford, believe they have won their case. And French officials say that their British counterparts expressed a preference for Daresbury at a meeting last week.

Rumours over the siting of the £175 million (US\$280 million) project intensified two weeks ago, after UK ministers met with the director-general of the Office of Science and Technology (OST) and the director of the Wellcome Trust.

Labour union officials at Daresbury say that Stephen Byers, the Secretary of State for Trade and Industry, has written to the trust saying the government is "minded" to site Diamond at Daresbury. The trust has also written to Daresbury's Member of Parliament (MP), Mike Hall, saying it has no strong views on where the source is sited.

Speculation on the decision has grown over the past three months (see *Nature* **401**, 197; 1999). Local unions, MPs and scientists have campaigned to site Diamond at Daresbury, which houses the current UK Synchrotron Radiation Source, rather than at the RAL.

Last Thursday, senior officials from the OST, the Wellcome Trust and the French science ministry met in Paris to begin negotiations that will lead to a memorandum of understanding over the Anglo-French project. France has agreed in principle to contribute FF350 million (US\$56 million) towards the construction costs of Diamond over seven years, and FF60 million to FF80 million a year to its operating costs (see *Nature* **400**, 489; 1999).

Michèle Leduc, an adviser at the ministry of research who attended the Paris meeting, said the British were likely to choose Daresbury — although a final decision awaits Byers' approval. "The British government

has been under tremendous pressure to site the machine at Daresbury," says Leduc.

Although the French government had lobbied for the RAL, because of its proximity to Oxford, Leduc said that science minister Claude Allègre still intends to cooperate on the project. "For the French this does not represent a break-up," says Leduc. "It is still the minister's intention to cooperate with the project scientifically."

Last Friday and Monday of this week the Wellcome Trust sent representatives to both sites to assess the 'risks' associated with the project. Trust officials refuse to say what these might be, but said the visit was "to draw information" for a meeting of the trust's governors on Wednesday (10 November). Further news is expected after the meeting.

Although the trust says that the visit does not suggest that Diamond would necessarily be sited at Daresbury, campaigners feel it is significant. Susan Smith, a union representative and scientist at Daresbury, said it was the first time the trust had officially been to Daresbury since the siting of the facility became an issue.

Whatever is decided, it seems that the Central Laboratory for the Research Councils (CLRC), which manages both sites, may emerge with its reputation tarnished. Its refusal to state a site preference, says Tony Bell of the Institute of Professionals, Managers and Specialists — the union that represents many research council scientists — has "destroyed morale amongst staff". He adds that expertise from Daresbury will be needed to build the site wherever it is located.

It also seems likely that the CLRC will not have sole responsibility for managing the site; concerns that it is not working well are reflected in a critical report from the National Audit Office.

Last week, science minister Lord Sainsbury confirmed that the government was considering options for the management of the facility. "We are pulling in views," says Sainsbury. CLRC's perspective is "one of the views".

Natasha Loder & Heather McCabe

Cambridge, UK, hopes MIT magic will work on joint institute

Washington

The Massachusetts Institute of Technology (MIT) has agreed to found an institute at Britain's University of Cambridge, funded by the UK government, intended to nurture industrial innovation.

Both universities envisage a broad-ranging partnership which, they say, will set the pace for global alliances between other major universities.

The UK government will provide £68 million (US\$110 million) over five years to launch the institute, which was announced in London on Monday by Gordon Brown, the UK Chancellor of the Exchequer. An additional £16 million from British industry will help the institute launch its first programmes.

These will include several MIT masters' courses in management and engineering at Cambridge. But the partnership is expected to expand to include exchange schemes for undergraduates and joint research in science and engineering, say senior officials at both universities.

Brown, who approached MIT after hearing that it was seeking a partnership with a European university, believes that the involvement of the famous US institution will help Britain address its long-standing difficulties in transferring technology from universities to industry.

"This is a path-breaking innovation," said Brown. "By choosing the United Kingdom as its European partner, MIT has recognized our country's strengths. The new institute will help us build on those by generating new ideas and by creating new business and jobs."

"We envisage a very broad engagement with Cambridge," says Larry Bacow, chancellor of MIT. "The intention is to stimulate innovation," he says, adding that this will involve basic research "in a number of areas, including neuroscience, biology, physics, materials and information technology".

UK politicians and industrialists have long envied the ability of top US research universities to stimulate innovation and economic growth. MIT is seen as the perfect model. A study estimated that its graduates have created corporations which would make up the twenty-fourth largest economy in the world (see *Nature* **386**, 103; 1997).

Bacow says that this will be MIT's only "institute-wide, broad-based partnership" with a European university. Colin Macilwain

Race is on to find successor to Varmus at NIH

Washington

The US government last week launched an expedited search for a new director of the National Institutes of Health (NIH), banking against long odds that it will be able to get a nominee confirmed before the presidential election next November.

Donna Shalala, the Secretary of Health and Human Services, charged the president of the National Academy of Sciences (NAS) with producing an unranked list of six to twelve names by mid-November. "We feel strongly that there should be a seamlessness in leadership at NIH," Shalala said in a statement.

The \$16 billion-a-year biomedical agency will be without a permanent director from the end of the year, when Harold Varmus leaves to become president of the Memorial Sloan-Kettering Cancer Center in New York City (see *Nature* **401**, 627; 1999). Deputy director Ruth Kirschstein will become acting director.

Bruce Alberts, the NAS president, says the Clinton administration hopes to see the Senate act on a nominee next spring. The Senate must confirm NIH directors but, before the administration sends a nomination to the Senate, it conducts a lengthy background check and political vetting. If a nomination does not arrive in the Senate by late winter or early spring, observers say, the odds of a confirmation become vanishingly small, because of the length of the confirmation process and the vagaries of election year politics.

The Senate tends to slow nominations as elections approach, to give the incoming administration a chance to make important



Fauci: known for his work on AIDS crisis.



Klausner: tough and able administrator.

appointments. The NIH directorship is also a key target for conservative Republicans — especially those facing re-election in anti-abortion states. Some are said to be planning to make the nominee's position on stem-cell research a litmus test.

Getting the nomination through "will be a challenge", said one Republican aide to the Senate Health and Education Committee. "The position offers so many opportunities for people to take symbolic stands on issues in an election year."

The insecurity of an appointment made by a waning administration may deter potential candidates, who could find themselves out of work after a change of administration. Elizabeth Marincola, executive director of the American Society for Cell Biology, says the conventional wisdom is that "nobody in their right mind would take a job under those circumstances". The solution, she says, is a candidate with "such incredible credentials that he would be almost immune to a turnover in administration. It's completely dependent on

the quality of person they recruit."

Alberts has contacted more than 1,000 scientists and scientific organizations. "Shalala would like her nominee to be a widely recognized, distinguished scientist," he wrote in a letter requesting the names of suggested candidates by last Monday (8 November).

Alberts told *Nature* that Shalala did not insist that candidates should be physicians, although most NIH directors have been medical doctors. He said that an overt political affiliation would rule out a candidate.

Among the key names being mentioned as potential nominees in Washington biomedical research circles are Anthony Fauci, director of the National Institute of Allergy and Infectious Diseases, and Richard Klausner, director of the National Cancer Institute.

Fauci, 58, who has been at the NIH since 1968, developed a national profile as the top official in the agency battling against the AIDS crisis. "He has been on every talk show 15 times. And he has handled a complex political and scientific challenge," says Mike Stephens, a lobbyist for the Federation of American Societies for Experimental Biology.

Fauci was offered — but declined — the NIH directorship by President George Bush. This could work for him in the Republican-controlled Senate, but might deter the Clinton administration.

Klausner, 48, a Varmus protégé, has been at NIH since 1979. He earned a reputation as an able administrator by heading a review of the NIH's intramural programme which resulted in its modernization, but he is not as well known as Fauci. His success in leading the largest NIH institute since 1995 — and persuading Congress of the importance of cancer research — are clear strengths.

Francis Collins, 49, director of the National Human Genome Research Institute, is being named by some as another strong internal candidate.

Names of outside candidates being floated include Michael Brown, 58, and Joseph Goldstein, 59, of the University of Texas Southwestern Medical Center at Dallas. Together they won the 1985 Nobel Prize for Physiology or Medicine for the discovery of the low-density lipoprotein receptor, and how mutations in it cause familial hypercholesterolaemia.

Also being mentioned are J. Michael Bishop, 63, chancellor of the University of California at San Francisco, who shared the Nobel prize with Varmus in 1989 for the discovery of the cellular origin of retroviral oncogenes; and Eric Lander, 42, who directs the Whitehead Institute/MIT Center for Genome Research in Cambridge, Massachusetts.

Herb Pardes, 65, vice-president for health sciences and dean of the faculty of medicine at Columbia University, New York, is considered a contender by some. **Meredith Wadman**

Surprise boost for stem-cell work

Washington

Strom Thurmond (Republican, South Carolina), a powerful conservative and the senior member of the US Senate, has testified in favour of allowing stem-cell research at the National Institutes of Health (NIH), increasing the chances that the Congress will allow public funding of research on stem cells derived from human embryos.

Thurmond's testimony showed that some anti-abortion politicians in the United States, who might have opposed such research, will support NIH funding for stem-cell research on the basis of its potential health benefits.

"The NIH must be actively involved in stem-cell research," the 96-year-old senator told a 4 November hearing of the appropriations subcommittee for labour, health and education, chaired by Arlen Specter (Republican, Pennsylvania). "NIH financial support of this research will allow us to move toward clinical applications."

As a measure of Thurmond's influence, Specter estimated at the hearing that Thurmond's previous support for federal funding of fetal-tissue research had doubled the number of Senate votes in its favour, from about 40 to 80.

Supporters of stem-cell research hope that Thurmond's testimony will encourage other conservatives in the House of Representatives and the Senate to support federal funding. Senator Trent Lott (Republican, Mississippi), the leader of the Senate, has promised that it will consider legislation on the matter in February.

At the same hearing, Congressman Jay Dickey (Republican, Arkansas) reiterated his opposition to stem-cell research and to the NIH's position, which is that it will not pay for extraction of stem cells from embryos, but will pay for research using cells that have been extracted by private organizations. Dickey predicts that the NIH's position will be challenged in the courts. **Colin Macilwain**

University debate boils over in Australia

Sydney

A radical plan by Australian education minister David Kemp to regulate the country's universities has blown up in the government's face, leaving the university system in a policy vacuum but turning its future into a highly charged political issue.

After Kemp denied in Parliament that he was contemplating major changes to policy, Labour's spokesman on education, Michael Lee, and opposition leader Kim Beazley dramatically revealed the minister's contradictions in a leaked submission to the Cabinet.

In it, Kemp disclosed: "Universities are currently in a difficult financial position ... Already, eight institutions appear to be operating at a deficit and some regional campuses are at risk."

The paper admitted that higher education "will remain a contentious issue for the government through this term [three years from October 1998]. Higher student-staff ratios, less frequent lecture and tutorial contact, ... outdated technology and gaps in the key areas of professional preparation are fuelling a perception of declining quality."

Kemp attributed this situation to "the highly regulated nature of the existing system", which he wanted to replace with a "demand-driven system". His solution was to deregulate student fees and numbers and open universities to "market forces". He wanted to give students portable vouchers, and loans at market rates, to insist on workplace reforms for staff and to gear courses and research towards commercialization.

Several senior academics had already advocated some of Kemp's proposals. Although these were bitterly opposed by staff and student bodies, all groups condemned the plan's secrecy. Lee estimated that the proposed loans scheme could take a graduate's debt to A\$100,000 (US\$64,000), and John Niland, president of the Australian Vice-Chancellors' Committee, said there were "a number of things worthy of consideration, but real rates of interest [on loans] is not one of them".

A public outcry and panic among the government's backbenchers over the threat to smaller universities triggered a hasty reversal of Kemp's plan by prime minister John Howard and his Cabinet. One plank survived, a controversial form of salary bargaining that caused a serious dispute with academic staff.

In a Green Paper on research, Kemp had revealed that he favoured reform based on restructuring rather than restoration of cuts (see *Nature* 400, 703; 1999). But while collecting reactions to this document, he decided that universities should be financed predominantly by students through portable vouchers, as recommended by a



Twist and turn: the University of Sydney must try to squeeze in more students with less funding.

review of higher education by Roderick West (see *Nature* 392, 854; 1998). Kemp appeared to have shelved its controversial proposals before the October 1998 election.

Some leaders welcome the opportunity to push alternative policies. Gavin Brown, vice-chancellor of the University of Sydney,

sees "a bright side", saying the leak brings "the urgency of the debate on higher education to the broader Australian public for the first time".

Brown favours deregulation, but believes it has been derailed by Howard's refusal to allow universities to set their own fees. He accuses Kemp of "engaging in an astonishing exercise in systematic alienation".

Brown says Australian funding is "unstable by any standards", being a sixth of that per student in the United States. He wants Australia to adopt Ireland's incentive for increasing external income through a government scheme of matching grants.

Alan Gilbert, vice-chancellor of the University of Melbourne, says "everybody is a loser". Accusing both Howard and Beazley of having "rushed to judgement, ruling out one policy option after another", Gilbert claims they "may jointly have condemned Australian higher education policy to long-term paralysis".

Mary O'Kane, vice-chancellor of the University of Adelaide, foresees a "dangerous year ahead", when the government "doesn't know what it is doing and says it has maintained funding when it has demonstrably cut it".

Peter Pockley

German scientists must repay fees

Munich

The German Administrative Court — the country's highest appeals body — has upheld a ruling that two retired medical professors from the University of Cologne must repay several million deutschmarks of fees for extra income gained from private diagnostics work.

With the agreement of both the university and the state science ministry, Gerhard Pulverer, then head of its Institute of Medical Microbiology, and Hans Eggers, then head of the Institute of Virology, had employed technical staff since the early 1970s with revenues from their private diagnostic work carried out for local hospitals and physicians in fields such as diphtheria and HIV.

In Germany, the additional income of university professors is normally subject to a fee for using state-funded equipment and staff, known as *Nutzungsentgelt*, of up to 50 per cent before tax. But Pulverer and Eggers had arranged with the science ministry of the *Land* (state) of North Rhine-Westphalia to pay only 17.5 per cent *Nutzungsentgelt*, in exchange for paying the salaries of additional technicians out of their own pockets.

These technicians, between 20 and 40



in number, carried out the laboratory work for the professors' private diagnostics. Apart from this, they were fully involved in the — previously severely understaffed — institutes' regular scientific work.

Jens Peter Meincke, the university's rector, confirms that the scheme benefited all parties, and that clinical diagnostics had gained significantly as a result.

But the court argued that the state-funded institute staff had been partly used

for the institute heads' extra work. This, it ruled, classified the institutes as *Mischbetriebe* (mixed business), which, according to a 1982 decree, should have ruled out the reduction of the *Nutzungsentgelt*. Regular institute work and the extra occupations of technical staff, it argued, cannot be financially balanced.

Pulverer and Eggers had maintained their scheme until 1987. They claim to have received repeated verbal encouragement to do so from the university administration and ministry officials.

They now feel the victims of the unclear legal situation after 1982. Pulverer blames the ministry for its wrong advice. "The ministry's legal advisers let us walk straight into the trap," he says.

He disagrees with one judge's view that, being university professors, they should have been intelligent enough to realize they were on thin ice. "We are medics and not lawyers," he says.

Including interest, Pulverer and Eggers will be hit with bills of about DM5 and DM2.5 million [US\$2.6 and 1.3 million], respectively. "My entire savings will not be sufficient," Pulverer says, "I will also have to sell my art collection."

Quirin Schiermeier

Israel shuts four 'incubators' for high-tech companies

Jerusalem

Four of Israel's 16 publicly supported 'incubators' for high-tech entrepreneurship are to close next year, says the office of Israel's chief scientist. The incubators were set up in 1991, originally to help immigrant scientists from the former Soviet Union develop their research.

Dafna Zamir, assistant to chief scientist Orna Berry, says that the closures are intended to make the programme more efficient, and that there will be no cuts in the number of projects accepted. No decision has been taken on which incubators will close, says Zamir. But the main criterion will be the number of projects that they have been able to attract.

The decision to close the incubators follows a fierce debate in the country's newspapers. Critics, including economists and some entrepreneurs, argue that the programme is a waste of money that has failed to produce profitable companies. But the programme's director, Rina Pridor, denies this and has fought successfully to preserve the current level of funding in the face of a move by treasury officials to make cuts.

The programme provides US\$300,000 in seed funds over a two-year period, as well as

managerial and administrative assistance. According to the latest figures, which cover the period up to June 1999, more than half of the projects that began in the incubators have continued under their own steam after leaving the framework, most with private funds.

There has been \$240 million of private investment in incubator projects since the programme began, notes Pridor, while the government has invested \$180 million.

One of the criticisms of the programme is that some of the incubators are in outlying areas rather than big cities. This was intended to attract high-tech industry away from the crowded central region and to make the incubators available to immigrants housed in these areas during a period when Israel had to absorb more than half a million immigrants in three years. Zamir says that geography will not be a consideration in deciding which ones to close.

Elie Engender, the director of the Ofek La'Oleh incubator, located in a small town near Haifa, says he's not worried, because he has no shortage of projects. He says the current grants make it difficult for incubator start-up companies to pay their researchers decent salaries.

Haim Watzman

\$2m ransom sought for kidnapped ecologists

Moscow

The Polish government last week rejected demands from Chechen kidnappers for \$1 million each for the release of two Polish scientists taken captive in August in Dagestan, one of the Russian North Caucasian republics bordering Chechnia.

Zofia Fiszor-Malanowska of the Warsaw Ecology Institute of the Polish Academy of Sciences, and Ewa Marchwinska-Wyrwal of the Ecology Institute of Katowice, had travelled to Dagestan at the invitation of Rasul Magomedov of the Biological Resources Institute of the Russian Academy of Sciences in Makhachkala, the capital of Dagestan.

The two ecologists arrived at the beginning of August and stayed at Magomedov's house. Soon all of them, together with Aleksander Karamasov, another Dagestanian scientist, moved to the Gunib region for ecological research. After the expedition failed to return at the expected time, the police started an investigation, but found only the car the scientists had used.

In September, relatives of Magomedov managed to negotiate his release. But his colleagues were left in captivity. A handwritten



Danger zone: this science base on the Abramov glacier, at 4,000-metres altitude in the Pamir mountains of Kyrgyzstan, was burned by terrorists in September. The base was set up in the former Soviet republic by the Central Asia Hydrometeorological Institute more than 30 years ago.

note from the hostages was planted on the Polish embassy in Moscow last week. In the note, the hostages ask for help because of their deteriorating health (both women are approaching 60). They are being held in a concrete cellar in Urus-Martan, Chechnia.

"We have no other information about our two citizens except that both are alive," says a spokesman for the embassy. The Polish

authorities declared that they will not pay the ransom. A Polish diplomat plans to visit Georgia in a bid to negotiate for the release of the women.

Kidnapping appears to have become an increasing hazard for researchers working in the former Soviet Union, where the activity has become an important source of income for dissident groups.

Carl Levitin

Publishers map out a way forward in response to free online archives

Commercial publishers plan to link up their journals, stealing the thunder of free repositories for life-science papers — and protecting their revenues.

Paris

Plans by the US National Institutes of Health (NIH) and the European Molecular Biology Organization to create a free, global repository for primary literature in the life sciences have prompted the major scientific publishers to cooperate, in principle, in setting up an alternative.

At a meeting in Frankfurt last month, 300 publishing executives agreed to link references in the articles they publish to the source papers in their respective publications. Some see this as a challenge to initiatives to provide the primary literature free to all — such as the NIH's PubMed Central (PMC) (*Nature* **401**, 6 & 626; 1999).

If the deal materializes, it will create a web of journal titles owned by different publishers. Publishers argue that for the many scientists with online access to journals via subscriptions paid for by their institutions, the result would be indistinguishable from PMC, as it would allow them to move rapidly and seamlessly from a citation — say from a Medline search — to full text.

David Lipman, director of the National Center for Biotechnology Information at the NIH and one of the main architects of PMC, says he is pleased that PMC has prompted publishers to cooperate. But he argues that the move is “qualitatively different” from PMC, whose goal is “barrier-free access for all,” including scientists who lack good library resources and the public.

Drawing the battle lines

Not everyone agrees. Richard Roberts, the 1993 Nobel prizewinner in medicine, describes the publishers' action as a thinly veiled “declaration of war” on PMC, and a bid to defend the status quo and the profit margins of scientific publishers.

Lipman is convinced that PMC will proceed. Two major journals, the *Proceedings of the National Academy of Sciences* (PNAS) and *Molecular Biology of the Cell*, have agreed to contribute their content, as have half a dozen smaller journals. The *British Medical Journal* (BMJ) is also likely to join. The concept is popular at the scientific grassroots, and Lipman hopes this will put pressure on non-profit learned societies to take part.

Richard Smith, editor of the *BMJ*, agrees. “There is a fundamental conflict for learned societies. PMC forces us to ask ourselves: is

our mission to make a lot of money from publishing to underwrite our activities, or are we there to advance science and medicine?”

Such arguments cut little ice in Frankfurt. “Naive idealism,” says the chief executive officer of one publishing house. One strong sentiment at the meeting was that PMC is an unhealthy government intervention in the scientific publishing market.

Lipman contests this, arguing that the NIH will only help run the infrastructure of PMC, and that peer review and other services will be provided by whoever wishes to take part — some in the private sector are already taking up the challenge (see page 110). As such, he argues that it is an infrastructure for the research community.

The other concern at Frankfurt was that ventures like PMC would undermine the print revenues of journals, push up prices and reduce quality. “We want to make back material as freely available as we can,” says Ellis Rubinstein, editor of *Science*. “But if we give away everything from as recently as one week ago, then we create a problem for ourselves, as subs would decline and those who wanted print or other services would pay more; we would lose our economies of scale”.

Smith predicts that most journals will come to exist only in electronic databases, and that the few that survive will be those offering editorial content other than primary literature. Anticipating this, the *BMJ* publishing group is shifting the focus of its specialized journals in this direction. Economies will also be made by making the full versions of papers only available on the web.

Nick Cozzarelli, editor of *PNAS*, believes that his journal's decision to put material on PMC one month after publication will protect its subscription base. “Free access is clearly in the interests of science, and a risk worth taking,” he says.

Roberts, who is editor of *Nucleic Acids Research*, published by Oxford University Press, is in discussions with the publisher aimed at putting the journal on PMC early next year, with papers reaching the repository on the same day as print publication.

His proposed business model is based on increasing page charges and decreasing subscription costs. Page charges would cover editorial costs and allow papers to be made freely available. “You don't need to be a rock-



Free-for-all? The private sector is already exploring alternatives to traditional publishing.

et scientist to see the benefits for scientists.”

But Frank Gannon, executive director of the European Molecular Biology Organization and the driving force behind E-Biosci — the European counterpart to PMC (see *Nature* **401**, 413; 1999) — predicts that repositories will only be able to amass enough content by cooperating with commercial publishers. As well as accepting material from authors, E-Biosci would link to material held on publisher's websites.

This contrasts with PMC, which requires that material be held on its own website. This is unacceptable, says Rubinstein. “*Science* cannot participate as PMC now stands,” he says, “we cannot put versions of our own copy on other sites, as it causes confusion and prevents us from adding enhanced services”. Lipman says that “everything is open to discussion”.

Although the contents of E-Biosci and PMC will be linked, it is becoming clear that they will to some extent compete. Plans for a single international advisory board have been abandoned. Instead, PMC and E-Biosci will have their own advisory boards; the international board will be made up of representatives of each board.

The PMC advisory board will include representatives from participating journals, librarians and the public, says Lipman, who approves of competition with E-Biosci. “I think it is a great idea,” says Cozzarelli, who predicts that the proposals will allay fears that the international board might be bureaucratic and dominated by US interests.

The NIH has asked the US National Academy of Sciences to help appoint the board, in what may be a first step towards transferring the lead role in PMC from the NIH to the academy.

Declan Butler

Brussels research chief seeks action on infrastructure funds

Barcelona European research commissioner Philippe Busquin acknowledged last week that current rules preventing the commission from providing infrastructure funding for life sciences research “interfere with research work in Europe”.

Busquin said that the commission was working towards finding a solution to the problem of funding the research centres affected. But he added that a long-term solution would have to await “a political decision of the 15 European Union countries”. Busquin was speaking in Lisbon after a meeting with the Portuguese research minister, José Mariano Gago.

Busquin's remarks followed the publication last week of a letter from 50 leading European biologists warning that the rules on infrastructure funding risk imposing significant damage on Europe's research base in the life sciences (see *Nature* **402**, 3–4 & 12; 1999).

He and Gago, the next chairman of the European Union research ministers, agreed that it was urgent to begin discussion of a ‘unified research policy’ within the union.

Gago said the rules limiting infrastructure funding were a symptom of a lack of coordination in research funding that underlined the need for a unified policy.

Panel sheds light on security at US weapons laboratories

San Diego An advisory committee to the University of California (UC) has found “inadequate concern for security” at two national weapons laboratories, but also believes that improvements show that science and security can co-exist. The conclusions come in a report issued last week by the UC President's Council Security Subcommittee, appointed in June to review the UC-managed Los Alamos National Laboratory in New Mexico and the Lawrence Livermore National Laboratory in California after allegations of Chinese spying.

The committee found that visiting foreign scientists are important for scientific exchange and can be monitored with current procedures, that classified computer systems are not ‘user friendly’ — prompting scientists to break rules and use unclassified computers — and expressed deep concerns about ‘false positives’ from lie-detector tests.

The eight-member committee is chaired by John F. Ahearne, a former chairman of the US Nuclear Regulatory Commission.

US action on health needs of minorities

Washington Bills have been introduced in the House of Representatives and the Senate that would upgrade the Office of Research on Minority Health at the National Institutes of Health (NIH) to the status of an NIH centre, with a view to placing more emphasis in US biomedical research on health issues affecting racial minorities.

Senator Edward Kennedy (Democrat, Massachusetts) and two Democratic senators for Hawaii introduced the Senate bill. In the House, Jesse Jackson (Democrat, Illinois) co-sponsored a companion bill that has bipartisan support. The legislation is more limited in scope than a previous proposal of Jackson's, which had been opposed by Harold Varmus, the outgoing director of the NIH. Supporters of the legislation say that Varmus now backs it, but he was unavailable for comment.

Japanese grants will unite industry and academics

Tokyo Japan's Ministry of International Trade and Industry (MITI) and Ministry of Education, Science, Sports and Culture (Monbusho) have created a research grant scheme to



Why get a new one even when the old one still works?

You are working in fields like gene therapy or genetic vaccination. You know the challenges of targeted and efficient gene delivery.

Viral vectors deliver genes efficiently, but can provoke strong immune responses and adverse reactions, which can be life-threatening. Plasmid vectors have no tissue selectivity. Moreover, 50% - 90% of the plasmid is made up of undesired sequences, such as antibiotic resistance markers and ori-sites. Many of these bacterial sequences contain immunostimulatory sequences (CpG motifs) that impair gene expression.

For the scientific success of your project, and also for the regulatory approval of your gene therapy, these are major problems.

Now you have the solution: the best of both worlds.

**A vector that is as smart as a virus,
yet completely safe and defined.**

MOLOGEN

www.mologen.com

Tel + 49 30 8417 88 17

Fax + 49 - 30 - 8417 88 50

midge@mologen.com

finance collaboration between industry and the academic community. It is the first grant scheme to be set up by different ministries.

The two ministries have asked for ¥9 billion (US\$85 million) from this year's second supplementary budget. They plan to fund 10–15 projects, in areas such as biotechnology, information technology and the environment, before the end of the fiscal year. MITI and Monbusho have promoted links between industry and academics since 1997, but bureaucratic barriers have made it difficult for collaborations to thrive.

Call to boost research on climate change

Munich Political discussion on the mechanisms and impact of global climate change are "inadequately backed by scientific research", the German Advisory Council on Global Change (WBGU) said last week. The statement coincided with the meeting in Bonn of the parties to the 1997 Kyoto protocol on curbing greenhouse-gas emissions.

The WBGU, an independent panel of scientists, said that the Bonn conference "missed an important goal", because the latest findings on the mechanisms of climate change "remained unconsidered". The panel warned, for example, that the predicted

change of some of the Earth's most important forest ecosystems from carbon sinks to carbon sources would probably accelerate global warming.

MIT appeals for \$1.5bn to step up research effort

Boston The Massachusetts Institute of Technology (MIT) is campaigning to raise \$1.5 billion to support research and education, and to pay for the construction and renovation of laboratories and buildings. Areas targeted include bioengineering, medicine, neuroscience and environmental studies.

The money will also be used to attract high level faculty members through higher salaries and increased financial aid. "It is our challenge to solve the next generation of great problems," says MIT president Charles Vest. In the 1950s, MIT bet on the potential of basic biology and "won," he added. "Today we have the knowledge and the courage to believe we can do it again" in other areas, he said.

Online lab supplies service set for launch

London A venture capital company, Atlas Venture, is to invest US\$6 million in an online business aimed at capturing a share of the

European life-sciences market. Universities, pharmaceutical companies and biotech companies in Europe spend \$5.3 billion a year on basic laboratory supplies.

eLabsEurope wants to increase online purchasing among researchers, saying that this will reduce the cost of supplies bought at high volume. The online marketplace — where labs will be able to search for and purchase supplies — goes live next March and expects to carry over 100,000 products.

Baumert wins inaugural European research prize

London The European Science Foundation (ESF) and the Latsis Foundation are to present a prominent researcher in science education with the first European Latsis prize for his work on education and human development.

Jürgen Baumert (below), of the Max Planck Institute for Human Development in Berlin, will receive the prize, of SFr100,000 (US\$64,000), at the ESF's assembly later this month. Baumert has studied the achievements of maths and science students in 30 countries.



MIDGE and MOLOGEN are registered trademarks of MOLOGEN GmbH

Safer and Smarter

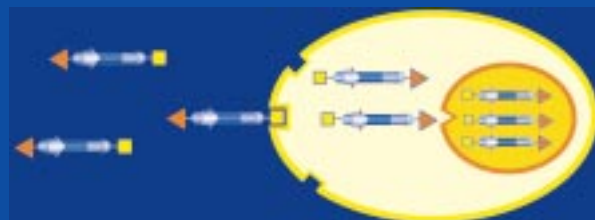
MIDGE vectors contain only the sequence information needed for the medical objective of gene transfer. Thus, the MIDGE vector in itself has no immunogenic potential and delivers only the gene you want. This is a strong argument for attaining regulatory approval.

MIDGES have been shown to have a considerably higher transfection and expression rate than plasmids.

We produce **MIDGE** vectors in GMP-quality according to your requirements from lab to commercial scale.

Do not miss the new generation of transfection vectors. Contact us today.

MIDGE allows for the covalent attachment of tissue specific ligands to deliver the vector only to target cells. And by using nuclear localization sequences nuclear entry is improved, resulting in high expression levels.



Survey confirms fears about licensing of genetic tests

Sir — S. M. Thomas *et al.*, in their surveys of patents for genetic diseases^{1,2}, expressed concern about monopolization by companies of “an entire gene and its mutations for all diagnostic and therapeutic purposes”¹. We asked a sample of owners of genetic testing patents about their licensing activities, and here we report that these patents are being exclusively licensed, permitting the monopolization of clinical testing services as Thomas *et al.* predicted. This result heightens our concerns that these patents may raise costs, curtail access to testing and limit clinical observation, which is fundamental to the practice of medicine^{3,4}.

We identified 37 US patents issued in 1991–97 that broadly cover the diagnosis of human genetic disorders. We excluded two patents held by non-US institutions and two that had expired. These 33 patents were assigned to 19 institutions, 16 public (such as universities) and three private companies. These patents cover tests for neurological diseases (13), cardiovascular diseases (6), metabolic disorders (6), cancers (5) and immunological disorders (3). Of the 33 patents in our analysis, 22 (67%) were funded at least in part by the US government, a percentage consistent with a previous sample⁵. With approval from our institutional review board, we called patent holders to ask them questions about licensing efforts, marketing of the patent or genetic test, known uses of the test, and enforcement of the patent. In summer 1998, we held 17 interviews covering 27 patents (response rate: 89% institutions; 82% patents). One respondent was an inventor and all others were involved in technology transfer or licensing at an assignee institution.

Our survey results are in Table 1. Nine of 14 patents (64%) issued more than two years before the date of our survey were licensed, in contrast to 5 of 13 (38%) of those issued more recently (one-sided Fisher's exact, $P = 0.17$). Tests for which respondents reported known clinical uses were more likely to have been licensed (OR = 14.7, $z = 2.3$, $P = 0.02$).

Critically, all 14 patent licences were exclusive. Although some authors have predicted that academic institutions would grant non-exclusive licences for patents⁶, these institutions do not have the resources to manage widespread licensing.

Exclusive rights to several of the patents in our sample are being used to monopolize testing services. Further, since our survey, US patents have been issued covering tests

Table 1 Patent survey results	N	%
Patents covered by survey	27	
<i>Licensing:</i>		
Exclusive licences granted	14	52
Non-exclusive licences granted	0	0
No known interest to take licence	9	33
Licensing in process	2	7
Will not be licensed	2	7
<i>Promotion and development of patented test*:</i>		
Active marketing effort	25	92
Not worth effort	1	4
Not being promoted, no reason given	1	4
Test undergoing commercial development	7	26
<i>Known uses of test*:</i>		
Clinical	9	33
Research	15	56
Don't know	8	30
<i>Require licence for research activities:</i>		
Yes, without exception	6	22
Yes, except academic researchers	6	22
Yes, royalty-free	3	11
No	12	44
<i>Enforcement of patents*:</i>		
Intend to enforce	18	67
No, too troublesome/expensive	9	33
Suspected infringement	5	18
Mailed notices of infringement	3	11
Brought lawsuit	1	4

*Totals exceed 27 because responses are non-exclusive.

for genes associated with breast and ovarian cancer, hereditary haemochromatosis, and spinocerebellar ataxia type 1. Each has been exclusively licensed to companies that are, to various degrees, restricting the performance of testing services by other laboratories. Not all patent holders are following this model. One such exception is the University of Michigan, which is broadly licensing its patent issued in 1998 covering the $\Delta F508$ allele that causes cystic fibrosis.

We believe that public institutions and the US National Institutes of Health should not grant exclusive licences for ‘upstream’ technologies and basic knowledge such as disease–gene associations, particularly when discovered using public funds. Such patents should be available to all, like the fundamental Cohen–Boyer patent on recombinant DNA (K. Ku, personal communication and ref. 7). Licensees should be permitted to perform patented tests non-exclusively and to sublicense others who wish to practice the patented methods. Collection of a reasonable royalty will ensure that the financial rewards and incentives of the patent system are maintained⁸.

Anna Schissel*, Jon F. Merz†, Mildred K. Cho‡
 *New York University School of Law, New York, New York 10012, USA

†Center for Bioethics, Department of Molecular and Cellular Engineering, and Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-3308, USA

‡Stanford University Center for Biomedical Ethics, Palo Alto, California 94305, USA

1. Thomas, S. M. *et al.* *Nature* **380**, 387–388 (1996).
2. Thomas, S. M. *et al.* *Nature* **388**, 709 (1997).
3. Cho, M. K. & Merz, J. F. *Nature* **390**, 221 (1997).
4. Merz, J. F. *Clin. Chem.* **45**, 324–330 (1999).
5. Merz, J. F. *et al.* *Mol. Diagnosis* **2**, 299–304 (1997).
6. White, T. J. in *The Genetic Frontier: Ethics, Law and Policy* (eds Frankel, M. S. & Teich, A. H.) 199–208 (AAAS, Washington DC, 1994).
7. Beardsley, T. *Nature* **311**, 3 (1984).
8. *Nature* **399**, 619 (1999).

Parents need catch-up courses to boost careers

Sir — You ask why there are so few women in science (Opinion, *Nature* **401**, 99 & 829; 1999). Discrimination and bias aside, the answer is because we are not men. Women can only follow a linear career path if they choose to have no children, or to put their children into full-time day care (if they can find it). Even highly motivated, intelligent women are often in conflict at work if, for example, they choose day care, or if they create large holes in their careers by taking time out to look after children.

At 37 years of age I will soon be finishing an MSc and starting an entry-level science job. I abandoned veterinary medicine early because of the incompatibility of an on-call schedule with family ties. The largest barrier that I currently face is how quickly a scientist's knowledge can become out of date. I do not feel that it is reasonable or possible for a professional to keep up to date while meeting the demands of a young family.

One solution would be to design a short programme, lasting less than a year, to bring parents up to date after they have had an extended leave (of greater than five years). This would at least qualify them for an entry-level position. I, by my own choice, have lost ten years on my ‘career path’. As the average career is between 40 and 50 years in length, I will have to be more than outstanding to ‘trickle up through the system’.

Wendy Powell

11 Whittaker Court, Guelph, Ontario, Canada N1C 1G1

Lacking cover

Sir — I was very surprised to see the advertisement for Dictagene on the back cover of the 7 October issue of *Nature*. Nudity has its place, as I am sure you will agree. It was totally unnecessary for a company to advertise its expression systems and custom-made proteins by displaying two nude human beings. What is this supposed to signify? I thought *Nature* was beyond such crass advertisements and I hope this is not the first of many such unpleasant sights.

Vani Kalyanaraman

Department of Anesthesiology, Washington University School of Medicine, St Louis, Missouri 63110, USA

Lessons in molecular gymnastics

Viewing the future of nanotechnology through rose-tinted spectacles.

Travels to the Nanoworld:
Miniature Machinery in Nature
and Technology

by Michael Gross

Perseus: 1999. 254 pp. \$25.95

Philip Ball

For his survey of nanoscale science and engineering, Michael Gross is so upbeat, so genial a guide that one feels churlish to dwell on the dark side. "Assuming that the peaceful use of nanotechnology flourishes and the military misuse will be suppressed efficiently," he breezes, appearing to imply that scientists have routinely managed this in the past. Meanwhile, China brandishes at troublesome Taiwan a wasp-sized spy helicopter measuring just $3 \times 5 \times 18$ millimetres. Not quite nano, but give them time.

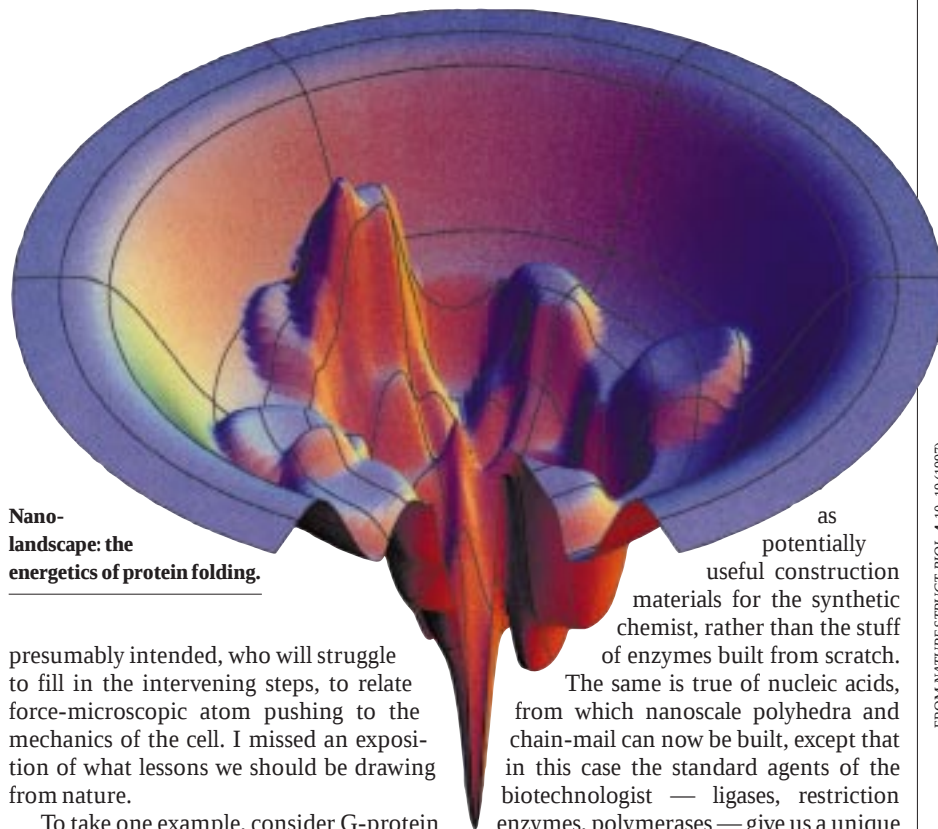
All technology is politicized, one point that has not eluded Eric Drexler in his campaign to advertise his vision of the nanotechnological future. Gross's book does not set out to address this dimension, and there is no absolute requirement that it should. Yet this is one illustration of the extent to which I missed critical analysis in an otherwise enjoyable tour of the nanoworld.

I was keen to read the book, because I imagined it would provide a choice selection of vignettes from the frontiers of chemical biology, and I was not disappointed. Here we have the molecular gymnastics of kinesin and myosin, the intricacies of chaperonin-assisted protein folding, the machinations of nitrogenase's inorganic heart.

From these biological nanomachines, we proceed to the factories: to cells that do clever things, such as orientate themselves to the geomagnetic field, register single-photon light sensitivity or engage in antibiotic chemical warfare.

Yet I was constantly left with the question: what does this tell the nanoengineers? Or perhaps more specifically, what does it tell them as opposed to the drug developers, the food scientists, the oncologists? The issue becomes still more pertinent in the chapter on biotechnology. I was fascinated to learn about the pressure-processed jams available in Japan, and glad to be updated on methods of cryopreservation; but where does the nano fit in? We are dealing here with weak interactions of the sort that are the bread and butter of supramolecular chemistry; but then, so was van der Waals, so (on occasion) was Faraday.

The biological world provides 'existence proofs' and endless inspiration to would-be nanotechnologists, but this is not enough. At least, I do not imagine it will be enough for the general readers for whom this book is



Nano-landscape: the energetics of protein folding.

presumably intended, who will struggle to fill in the intervening steps, to relate force-microscopic atom pushing to the mechanics of the cell. I missed an exposition of what lessons we should be drawing from nature.

To take one example, consider G-protein signalling. "Maybe future engineers can borrow a switch or two from here," suggests Gross; but, with an engineer's hat on, G-protein transduction looks absurdly cumbersome, "an intriguingly complex mesh of structural interdependencies and interactions", as the author admits. Would it be more fruitful to look, not at structural details, but at algorithmic questions such as receptor redundancy — should the artificial nose have one-analyte, one-receptor correspondence, or (as it appears) a specificity defined at the system-wide level?

Gross is more persuasive where he is most at home, with protein structure and design. What are the prospects for cracking the 'folding code', the relationship between sequence and structure? As Gross points out, for all the rapidity of structural solutions, we have data on only a pitiful fraction of this sophisticated language, and those are biased towards relatively small and easily crystallized proteins. The big, difficult cases might contain secondary motifs still unguessed.

With this in mind, the successes so far in protein design are impressive — although it seems significant that these have come largely from mutation of extant proteins rather than from *de novo* design, which does not (I am assured) yet achieve the same tightly packed, robust motifs as nature's constructions. Peptides are best regarded at present

as potentially useful construction materials for the synthetic chemist, rather than the stuff of enzymes built from scratch.

The same is true of nucleic acids, from which nanoscale polyhedra and chain-mail can now be built, except that in this case the standard agents of the biotechnologist — ligases, restriction enzymes, polymerases — give us a unique set of versatile tools.

Some of these developments are highlighted in a whistle-stop tour of supramolecular chemistry, which introduces all the usual suspects: crown ethers, rotaxanes, lock-and-key catalysts, Langmuir–Blodgett films, self-assembled monolayers. Fullerenes and carbon nanotubes get their due; however, here, minor errors leave me vaguely anxious about the material with which I am less familiar. Sumio Iijima, whose name is misspelt, might not appreciate being described as "infected" with "fullerene fever" (did anyone really call it that?), given his prior history of work on carbon microstructures; and Don Huffman's lab is a long way from Heidelberg.

There are other lapses in the chemical discussion. What, I wonder, is meant by calling the transition state "a futile arrangement of molecules"? I was puzzled about why gold chemistry was made to sound quite so odd ("it might almost be mistaken for a normal chemical element"), or why its practitioners should ever have been "suspected of secretly practising alchemy". And I thought the time long past when alchemy must be characterized as chemistry's "dubious heritage".

The style of writing is hard to place — I suspect the target audience is the legendary educated lay person, but the level is often close to a News and Views article in *Nature*.

FROM NATURE STRUCT. BIOL. 4, 10–19 (1997).

I sympathize with Gross about how easy it is to forget that words like 'moiety' are not common usage, or that there comes a point at which the accumulation of technical terms becomes gibberish to the non-scientist, no matter how carefully each term has been defined. Molecular science has a particularly difficult barrier to overcome, although Gross's humour and enthusiasm are considerable compensation.

But in the end, I'm left feeling that a topic as broad, indeed ill-defined, as nanotechnology needs a strong element of interpretation, evaluation and organization, before an outsider can see the what, the how and, most of all, the why. This book provides captivating snapshots, but sometimes lacks the narrative of a good movie, or book for that matter. ■

Philip Ball is a consultant editor at Nature.

On the same scale

Carbon Nanotubes and Related Structures: New Materials for the Twenty-first Century

by Peter J. F. Harris

Cambridge University Press, £50, \$80

Evolutionary play on maternal behaviour

Mother Nature: A History of Mothers, Infants and Natural Selection

by Sarah Blaffer Hrdy

Pantheon/Chatto & Windus: 1999.

725/448 pp. \$35/£20

Kristen Hawkes

This big book is about the complex trade-offs of motherhood, the ways those trade-offs have driven the evolution of our species, propelled cultural diversity and continue to shape our lives. Among other things, Sarah Hrdy considers the evolution of mammalian physiology, primate and non-primate female reproductive strategies and variation in human maternal behaviour as revealed in ethnography, history, psychology, sociology, journalism and fiction. Her startling and persuasive thesis, synthesizing a wide range of recent work, including her own, is that maternal strategies have played the star role in the evolution of our humanity.

She begins with the 'maternal instinct', contrasting observed maternal behaviour across many species with the myth of naturally self-sacrificing motherhood. Any organism that gives birth repeatedly over a lifetime must trade effort allocated to a particular egg, fetus or infant against effort allocated to another offspring — living or yet to be born. In a species where mothers can continue to invest in previous offspring long after bearing others, the choices are even

more complex. One new baby can put the future of many others at risk. Reproductive choices make a large difference to whether the mother has successful descendants. Hence, the molecular, physiological, emotional and even cognitive mechanisms underlying this decision have necessarily been shaped by natural selection.

Early on, Hrdy shows how much can be explained by viewing individuals as strategists whose capacities for assessing and responding to the circumstances of their lives are shaped by their evolutionary past. The example of the worker-queen switch in bees illustrates how large (and predictable) differences can result from differences in nurture among genetically equivalent individuals. The character and range of maternal

responses in any species are built from the materials at hand during its evolutionary history. In the case of humans, our general mammalian, and particularly primate, heritage is central to understanding the adaptive variety of maternal strategies.

More than 25 years ago, Hrdy hypothesized that the character and distribution of libidinous behaviour among female primates was actually maternally motivated. Because adult males can be a danger to the infants of other males, females reduce this threat to the survival of their offspring by spreading the possibility of paternity among many males. Here, reporting some of the mounting evidence favouring that hypothesis, she suggests that "pejorative-sounding words like 'promiscuous' only make sense



Nurturing the nature of art

Ana Maria Pacheco, the Brazilian-born sculptor, painter and printmaker, has been Associate Artist in residence at the National Gallery in London for the past two years.

Her remit is to create new work inspired directly by the gallery's collection of paintings.

The two examples of Pacheco's work shown above are the oil painting, *Queen of Sheba* and *King Solomon in the Garden of*

***Earthly Delights*, and a mother and child detail from a multi-piece figure sculpture, *Dark Night of the Soul*, both produced this year.**

An exhibition of Pacheco's work entitled "Ana Maria Pacheco in the National Gallery" is running at the gallery until 9 January 2000. A book of the same name (published by the National Gallery, £9.95, pbk) looks at her work at the gallery and her past career.

from the perspective of the males ... From the perspective of the female however her behavior is better understood as 'assiduously maternal'."

The middle part of the book explores the kinds of care that human mothers give their children. We share the legacy of breast-feeding with other mammals, and intense sociality with most primates. However, human mothers rely far more than other primates on others (allomothers) for direct contributions to the welfare of their offspring. The common assumption in popular scenarios of human evolution is that women get this help from their husbands. Hrdy reviews evidence that men *can* experience the same pleasure as women from care-giving intimacy, but there are good reasons why women are more likely to respond to infant signals first. And just as mothers can't escape quality/quantity trade-offs, males can't escape mating competition with other males. That may draw men away from care-giving, even if they are certain of paternity. "Earlier commentators failed to consider how unusual are the particular environmental and demographic conditions that make long-term monogamy advantageous for both sexes."

If fathers may not provide the needed help, there is another candidate available to human mothers, one perfectly suited for the job by experience and common interest: postmenopausal grandmothers. The unusually slow ageing that distinguishes us from other primates, as well as our need for allomothers, might have evolved when older females with declining fertility could help enough to have substantial effects on the maternal success of their younger kin.

Hrdy suggests that our reliance on allomothers lies behind another human difference. Other primates do not discriminate among their offspring by sex or state of health. In humans, however, contingent maternal commitment to children of different sex, age and condition is common. Hrdy suggests it arises from the more complex trade-offs that go with bearing a new baby while the previous child still needs care.

The remainder of this long middle section of the book discusses the extreme variations in human maternal commitment recorded ethnographically and especially historically. Hrdy reviews the frequency and distribution of infant abandonment and infanticide, the sex biases often observed in these practices, and the wet-nursing that swept eighteenth-century Europe. Throughout, she seeks the "shifting constraints and options" that mothers face and the composite of biological responses that can explain the nature of the variation in human mothering.

Contingent maternal commitment to offspring poses a special challenge for human infants. The final section of the book focuses on the offspring's side of the story,

beginning with the perils of surviving to birth and the insights into pregnancy that come from recognizing conflicting interests between mother and embryo.

Hrdy hypothesizes that, with new babies arriving while their older siblings were still in need of care, mothers who were more discriminating about the distribution of care fared better. Such discrimination gave survival advantages to infant features that increased the chance of maternal commitment, selecting for fat, playful babies, "the infantile equivalent of sex appeal".

In the final part of this last section, Hrdy tackles the arguments that have raged around the needs of infants and the contemporary aspirations of women. Her credentials as eminent scientist and teacher as well as passionate feminist and mother give her both authority and sympathy for arguments on many sides. She makes clear that merely understanding the conflicts of interest between mothers and infants, the sources of maternal ambivalence and the inevitability of trade-offs does not make them go away.

This book is a major contribution to the evolutionary biology of our species. By including some of her own intellectual and personal biography and attending to the history of ideas, Hrdy makes it also a contribution to the history and sociology of science. Anyone who thinks that working mothers and variable family arrangements are an unnatural recent novelty should read this book. Anyone interested in the causes (and consequences) of variation in women's behaviour, human sexuality or human evolution *must* read this book. It is superb human behavioural ecology. ■

Kristen Hawkes is in the Department of Anthropology, University of Utah, Salt Lake City, Utah 84112, USA.

..... Rocks of the ages

Reef Evolution

by Rachel Wood

Oxford University Press: 1999. 414 pp. £55

Robert W. Buddemeier

An ecological perspective on reefs and reef organisms might seem a shaky basis for a book dealing largely with ancient environments and extinct organisms from geological samples. However, the requirements and opportunities involved in competition for space on reefs, and how these have changed and developed over evolutionary time, is a central theme in Rachel Wood's book. In general, her gamble pays off. She provides a broader and more integrative level of biological insight into the development of both ancient and modern reefs than the palaeo-environmental and evolutionary discussions usually encountered in treatments of fossil reefs and reef organisms. The approach

carries with it a necessary focus on the biogenic, *in situ* framework as a defining feature of reefs, and this will not please fans of the mud-mound and the *Halimeda* bank. However, these 'reef-like' structures are also clearly and comprehensively, if somewhat peripherally, described and discussed.

Wood also gives an extensive and up-to-date treatment of the literature on modern coral reefs, and places contemporary issues of global change in the long-term context of reef geology and evolution (and vice versa). Here again, she takes a risk — emphasis on present knowledge in a rapidly changing field means that some points will date early in the useful life of the book. But in terms of advancing understanding and scientific synthesis, information that is partly out of date is almost always better than completely irrelevant information.

Those involved in reef studies tend to tout the multi- or interdisciplinary nature of the field, but all too often communication breaks down in a *mélange* of jargon and imperfectly defined or understood terms and concepts. Wood handles this problem outstandingly — a combination of clear links to a glossary, definitions in the text and simple avoidance of unnecessary and potentially confusing technical terms contribute to the lucid writing that characterizes this book. Carbonate geologists should take note: it is possible to write an informative book on the subject without once using the word 'facies'. My one criticism on clarity is that many of the illustrations are too small, and some would have benefited from more generous use of labels and markings.

Because of the interdisciplinary nature of the field, many of the comprehensive books on reefs are multi-authored collections. Even with a good plan, good editing and good summary chapters, such books don't give the consistency of style, organization or conceptual viewpoint that can make a single-author book more satisfying and informative. Wood intelligently presents a coherent philosophy, and does a creditable job of telling the reader on which topics her approach differs from that of others. Some of her positions will not be uniformly accepted, but they are taken with zest and imagination, and logically supported. I expect to enjoy the ensuing debates as much as I enjoyed the book.

I have a short book-shelf of key reference books that I consider particularly clear, comprehensive and accessibly organized, and which I use as my starting point for exploring unfamiliar topics and linkages. Pre-Cenozoic reef geology and evolution have not been satisfactorily covered in my mini-library to date; Wood's book fills most of the gap, and is going straight onto the 'short shelf'. ■

Robert W. Buddemeier is at the Kansas Geological Survey, University of Kansas, Lawrence, Kansas 66047, USA.

Growing old in Germany

The Berlin Aging Study: Aging from 70 to 100

edited by Paul B. Baltes and

Karl Ulrich Mayer

Cambridge University Press: 1999. 564 pp. £45, \$69.95

John Grimley Evans

Even if born of nothing more than convention, a multidisciplinary book on human ageing is to be welcomed in the modern academic world, where gulfs continue to widen between specialties and professions. Too many medical schools have libraries in which a restricted choice of journals ill-serves the conscientious geriatrician trying to maintain literacy in psychology, ethology or social policy. The reciprocal problem is at least as great for social scientists.

This book is the first phase of a study by the Berlin-Brandenburg Academy of Sciences of a group of old people living in what was West Berlin. The study covered four theoretical aspects: differential ageing; continuity versus discontinuity in ageing; plasticity and reserve capacity in old age; and ageing as systemic and interdisciplinary phenomena. Four highly distinguished research teams outline findings in the fields of medicine, psychiatry, psychology and socioeconomics. This is followed by more or less interdisciplinary analyses of topics such as sensory systems, competence in everyday skills and sources of well-being in the elderly group. The book was originally published in German but has been translated into limpid American.

The investigators were careful to nominate a representative sampling frame of 1,908 subjects, mainly using the city registry. However, the final response rate was 27 per cent, 516 subjects, not impressive by modern epidemiological standards. Intensive analysis of the selection stages suggested that only minor bias was introduced in the selection, but the reader will wonder to what extent so attenuated a sample from West Berlin can be generalized. The methods used, however, were rigorous and systematic, and the scholarship is formidable in breadth and perspicacity.

A central limitation of the survey, which the editors acknowledge, is its cross-sectional design. Differences between older and younger subjects may be due to ageing, but might also reflect selective survival or cohort effects. Some uncertainties will be resolved by the planned longitudinal follow-up, which I hope will be linked to a new cohort of younger subjects. This classic cross-sequential design is essential in attempting the fundamentally insoluble

resolution of age, period and cohort effects in gerontology.

The ambiguities inherent in interpreting cross-sectional studies are exemplified by the geriatrics team, which noted that, although the prevalence of coronary heart disease was linked to the conventional risk factors in the younger segment of their sample, this relationship was attenuated in the older subjects. They attribute this to selective survival. "In our opinion..." (a phrase one hears nowadays only from continental doctors; under the martial law of evidence-based medicine, if British doctors still have opinions they are no longer at liberty to express them) ... "in our opinion, our results ... should provide sufficient reason to prescribe lipid-lowering medication to an elderly patient with a fat-metabolism disorder not yet manifested in atherosclerosis". One need not be a health economist to worry about the implications as well as the logic of this statement.

The book is full of intriguing observations and hypotheses. The effects of socioeconomic disadvantage on health and well-being do not seem to carry through from working life into old age as markedly as they do in Britain and the United States. This is attributed both to a more realistic system of social welfare and to a failure of some of the rich to take out adequate insurance. That the poor should be protected against penury and the rich penalized for lack of foresight seems

a rational basis for an egalitarian society. The survey indicates that most old people in Berlin seem well adjusted to their state. Nonetheless, in extreme old age some reach the limits of their reserves and begin to show psychological traits indicative of critical levels of stress.

The value of this work would have been enhanced by a comparison with similar studies in other Western countries, allowing some of its speculations to be probed. Unfortunately, few other European nations have made a similar research commitment to the welfare of their ageing populations. I would also have liked to have seen an historian included in the investigators. Citizens of Berlin born between 1887 and 1922 lived through some interesting times. The suffering of the German people in both world wars was appalling, and some will have lived through the heroic 1945 defence of Berlin. Perhaps history was considered too sensitive a topic to be part of the study. A passing comment hints that some subjects were disinclined to be involved in politics in their old age because of past political attachments. A pity; the fruit that individual experiences bear in our later years is an intriguing area of gerontology.

John Grimley Evans is in the Division of Clinical Geratology, Nuffield Department of Clinical Medicine, University of Oxford, Radcliffe Infirmary, Oxford OX2 6HE, UK.



Holding a mirror up to da Vinci's science

Bill Gates still holds the record for the highest price paid for a printed book sold at an auction. In 1994 he bought Leonardo da Vinci's Codex Leicester — a 72-page study of the Earth's hydrodynamics which Leonardo wrote in mirror-image between 1506 and 1510 — for \$30.8 million.

Having been exhibited in Italy, France, Portugal and the United States, the Codex is now on display in Germany. The original folios are accompanied by banks of computers, where visitors can view the electronic versions of the pages, reverse the images to read the original

Latin text or the German translation. Scientists must regret, however, that, unlike the annotated electronic versions of important manuscripts of Galileo held by the Italian National Library of Florence (<http://www.imss.fi.it>), the electronic pages of the Codex are not available on the Internet.

The exhibition runs until 9 January 2000 at the Haus der Kunst in Munich. It then moves to the Museum der Dinge in the Martin Gropius Building in Berlin, where it will be shown from 30 January until 18 March 2000.

Dark on a light subject

Nineteenth-century physicists couldn't agree about the speed of light.

Brian Pippard

The six volumes of the English mathematician and physicist Lord Rayleigh's collected papers record a life's work unusually free from mistakes. One paper of 1881, however, is partly omitted as erroneous. This paper was the opening shot in a good-humoured conflict about the speed of light among leading physicists in Britain and America, with a little help from France — most is told as letters to *Nature*.

We must first go back 30 years, to 1851 and the determinations of the speed of light by Armand Fizeau and Jean Foucault in Paris. Fizeau used a rapidly spinning toothed wheel to chop light into short sections which were reflected from a mirror 8.6 km away; they took long enough to return so that, with the wheel spinning at certain speeds, they were obscured by a tooth. Foucault reflected light from a spinning mirror and thence to a distant fixed mirror, so that what returned found the spinning mirror had turned to displace the image.

What concerns us here is the repetition of Fizeau's measurement by James Young and George Forbes of Glasgow, who reported in 1881 that blue light travels 1.8% faster than red, and Rayleigh's comment on their work. Their result was more credible in Glasgow than elsewhere, for Lord Kelvin did not quite accept James Clerk Maxwell's theory of light as electromagnetic waves and was (rightly) satisfied that Young and Forbes were capable experimentalists. Elsewhere, however, there were sceptics, particularly among those astronomers who had never observed a celestial body displaying a succession of colours as it emerged from eclipse. Across the Atlantic, Albert Abraham Michelson, an old hand at the speed-of-light business, wrote that in his measurements the effect would have spread the image of the source slit into a spectrum.

Michelson had written within days of reading Rayleigh's critique of Young and Forbes, which raised the question of what speed is actually measured. In his *Theory of Sound* (1877) Rayleigh had shown that a limited train of waves, such as that produced by Fizeau's wheel, would travel in a dispersive medium at a speed, u , different from that of the individual wave crests, v . If the wavenumber ($2\pi/\text{wavelength}$) is k and the angular frequency is ω , then $v = \omega/k$ and $u = d\omega/dk$. The results of Young and Forbes would imply that v is 2.7% less than u , and this is inconsistent with the value of v that alone comes from measurement of the aber-

ration of stars. All other determinations, said Rayleigh, are of u .

This remark started a new hare when Rayleigh was asked to explain his belief that Foucault's method measured the speed of light as u . He obliged, but had to admit that he had overlooked a complication. When light is reflected from a spinning mirror, one side is moving towards the light source, and the other side away. There is therefore a different Doppler shift from the two sides, so that the outgoing wavelength varies across the width of the mirror, and in a dispersive medium the two sides of the wavefront move at different speeds, and the wavefront tilts during its long journey. This results in an extra displacement of the image, so that it is neither v nor u that is measured, but v^2/u . He added that a convex lens in front of the fixed mirror would invert the wavefront and allow the extra tilt to be cancelled, so that now u would be measured after all.

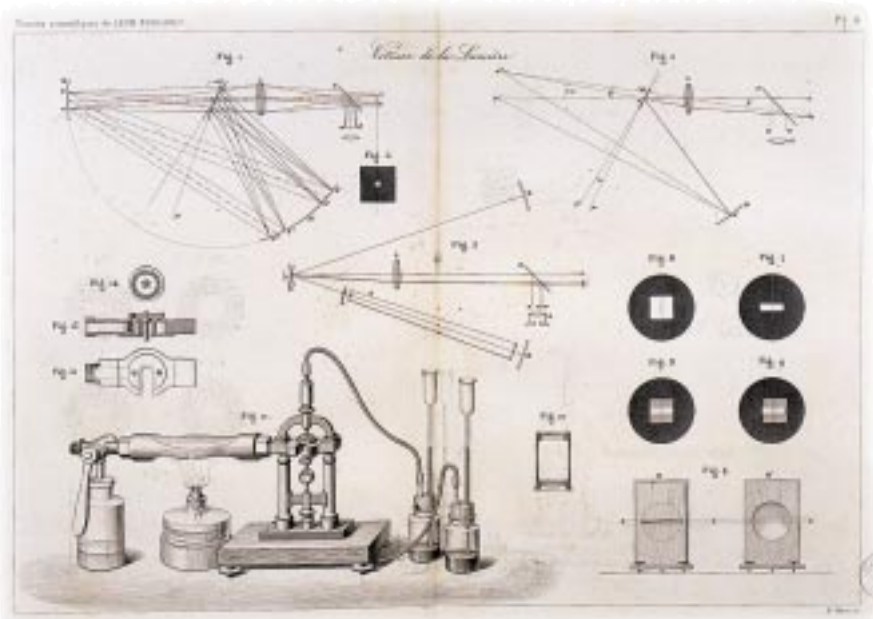
Nothing much happened for five years, until Arthur Schuster wrote from Manchester to dispute George Gouy's opinion, "given

with insufficient reason", that u is correct, and to question Rayleigh's analysis of the effect of a convex lens; it wouldn't cancel the Doppler effect, but enhance it, to give the measured speed as $v^2/(2v-u)$. Schuster says that Rayleigh concurs, but now Willard Gibbs of Yale enters the ring, moved by Schuster's letter to dispose of the whole matter in magisterial fashion: "I am not aware that attention has been called to the important fact, that while the individual wave rotates the wavefront of the group remains unchanged... To get a picture of the phenomenon, we may imagine that we are able to see a few inches of the top of a moving carriage wheel. The individual spokes rotate, while the group maintains a vertical direction." The right answer, after all, is u , and Gibbs fires a Parthian shot at Schuster — if he hadn't made a false assumption he too would have got u .

This closes the discussion, with honours to America — Michelson and Gibbs spot on; and the British team of Young and Forbes, Kelvin, Rayleigh and Schuster all a little off-beam. But what went wrong with the Glasgow measurements? It is hard to fault the very full account they gave in 1882 of their prolonged and scrupulous research. All the same, it was vitiated by an unexplained effect and, like many another honest error, has been consigned to oblivion. ■

Brian Pippard is at the Cavendish Laboratory, Department of Physics, University of Cambridge, Madingley Road, Cambridge CB3 0HE, UK.

Young and Forbes reported that blue light travels 1.8% faster than red.



It's all done with mirrors: Foucault's apparatus for measuring the speed of light.

Homo sapiens declared extinct

Yes, human beings have finally gone, but the 24-hour global party continues.

Bruce Sterling

AD 2380: After a painstaking ten-year search, from the Tibetan highlands to the Brazilian rainforests, it's official — there are no more human beings.

"I suppose I have to consider this a personal setback," said anthropologist Dr Marcia Raymo, of the Institute for Retrograde Study in Berlin. "Of course we still have human tissue in the lab, and we could clone as many specimens of *Homo sapiens* as we like. But that species was always known primarily for its unique cultural activity."

"I can't understand what the fuss is about," declared Rita "Cuddles" Srinivasan, actress, sex symbol and computer peripheral. "Artificial Intelligences love to embody themselves in human forms like mine, to wallow in sex and eating. I'm good for oodles of human stuff, scratching, sleeping, sneezing, you can name it. As long as AIs honour their origins, you'll see plenty of disembodied intelligences slumming around in human forms. That's where all the fun is, I promise — trust me."

The actress's current AI sponsor further remarked via wireless telepathy that Miss Srinivasan's occasional extra arms or heads should be seen as a sign of "creative brio"; and not as a violation of "some obsolete, supposedly standard human form".

A worldwide survey of skull contents in April 2379 revealed no living citizen with less than 35 per cent cultured gelbrain. "That pretty well kicks it in the head for me,"

declared statistician Piers Euler, the front identity for a collaborative group-mind of mathematicians at the Bourbaki Academy in Paris. "I don't see how you can declare any entity 'human' when their brain is a gelatin lattice, and every cell of their body contains extensive extra strands of industrial-strength DNA. Not only is humanity extinct but, strictly speaking, pretty much everyone alive today should be classified as a unique, post-natural, one-of-a-kind species."

"I was born human," admitted 380-year-old classical musician Soon Yi, speaking from his support vat in Shanghai. "I grew up as a human being. It seemed quite natural at the time. For hundreds of years on the state-supported concert circuit, I promoted myself as a 'humanist', supporting and promoting human high culture. But at this point, I should be honest: that was always my stage pretence. Let's face it: gelbrain is vastly better stuff than those grey, greasy, catch-as-catch-can human neurons. You can't become a serious professional artiste while using nothing but all-natural animal tissue in your head. It's just absurd!"

Beasts, birds, butterflies, even the very rocks and rivers must be rejoicing to see humans gone.

Gently fanning his wizened tissues with warm currents of support fluid, the grand old man of music continued: "Wolfgang Mozart was a very dull creature by our modern standards but, thanks to gelbrain, I can still find ways to pump life into his primitive compositions. I also persist in finding Bach worthwhile, even in today's ultracivilized milieu, where individual consciousness and creative subjectivity tend to be rather rare, or absent entirely."

Posthumanity's most scientifically advanced group, the pioneer Blood Bathers in their vast crystalline castles in the Oort Cloud, could not be reached for comment.

"Why trouble the highly prestigious Blood Bathers with some trifling development here on distant Earth?" demanded President Arno Hopmeier of the World Anti-subjectivist Council. "The Blood Bathers are busily researching novel realms of complex organization far beyond mere 'intelligence'. We should feel extremely honoured that they still bother to share their lab results with creatures like us. It would only annoy Their Skinless Eminences if we ask them to fret over some defunct race of featherless bipeds."

A Circumsolar Day of Mourning has been declared to commemorate the official extinction of humanity, but it is widely believed that bursts of wild public enthusiasm will mar the funereal proceedings.

"When you sum them up," mused Orbital Entity Ankh/Ghiih/9819, "it's hard to perceive any tragedy in this long-awaited event. Beasts, birds, butterflies, even the very rocks and rivers must be rejoicing to see humans finally gone. We should try to be adult about this: we should take a deep breath, turn our face to the light of the future, and get on with the business of living."

"Since I've been asked to offer an epitaph," the highly distributed poetware continued, "I believe that we should rearrange the Great Wall of China to spell out (in Chinese of course, since most of them were always Chinese) — 'THEY WERE VERY, VERY CURIOUS, BUT NOT AT ALL FAR-SIGHTED.'"

"This historical moment is a serious occasion that requires a sense of public dignity. My dog, for instance, says he'll truly miss humanity. But then again, my dog says a lot of things."

Bruce Sterling (<http://www.well.com/conf/mirrorshades/>) is the author of *Schismatrix* and many other novels and stories; the non-fiction work *The Hacker Crackdown*; co-author (with William Gibson) of *The Difference Engine*; and editor of *Mirrorshades: The Cyberpunk Anthology*.



OLIVER BURSTON

Survival of the weakest

Erik Asphaug

Many asteroids have a porous structure, which helps them absorb energy from collisions without being smashed to bits. This shows how planets might have formed, but could be bad news should an asteroid ever threaten Earth.

There are tens of millions of asteroids in the Solar System; a few are hundreds of kilometres across, but most are mountainous lumps of rocky material, orbiting mainly between Mars and Jupiter. Some asteroids swing closer to Earth than any other planetary object, yet we know less about them than we did about the Moon 50 years ago. Asteroid geology is thus a young and exciting science, as shown by the provocative hypothesis put forward by Housen and colleagues on page 155 of this issue¹ to account for the giant craters seen on some asteroids. Their idea — that many asteroid craters are formed by material being crushed rather than explosively ejected — is surprising, but it has the potential to resolve a long-standing question about how planets form. Housen and colleagues' hypothesis also spells a cautionary tale, because plans to disrupt or nudge a threatening asteroid out of harm's way might falter if it can readily absorb impact energy.

The dominant geological process affecting asteroids is collision with other asteroids. Most of these collisions produce craters, formed when material is blown explosively from the impact site. The most violent collisions are catastrophic and give birth to families of smaller asteroids. At gentler speeds, colliding asteroids can merge, and this is important for understanding how the planets might have formed from the swarm of planetesimals (proto-asteroids) in the early Solar System. Because the weak gravity associated with small asteroids strongly discourages mergers, modellers have no easy time getting asteroids to stick together under realistic conditions. Seismic waves from a collision that would hardly be noticed on Earth can move mountains on a low-gravity asteroid, excavating a multi-kilometre crater or breaking up the asteroid into pieces. Because cratering (let alone a catastrophic collision) eats away at an asteroid's mass, researchers have long wondered how asteroids survive, and how planetesimals could have accumulated into planets.

One explanation was that internal strength must do the work where self-gravity fails to hold an asteroid together; that is, asteroids are monoliths resistant to fracture. But this idea could barely explain their survival, let alone their accumulation, and has begun to fall out of favour. All known asteroid densities are quite low, suggesting they

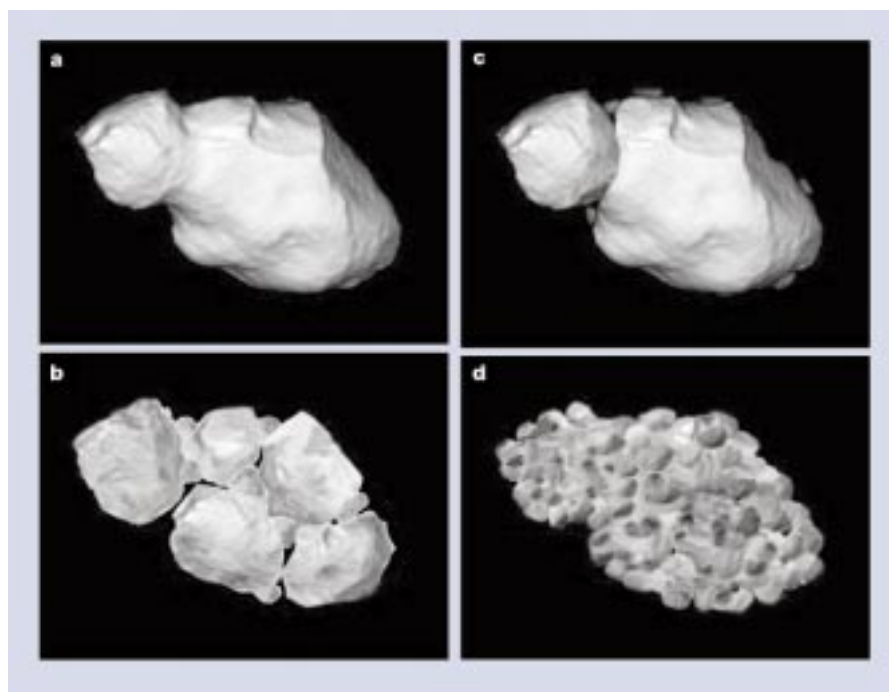


Figure 1 What are asteroids made of? The shapes of asteroids are being determined with ever-increasing accuracy, and their surfaces imaged in ever-increasing detail, but scientists can only guess at their interiors. To illustrate the point, a high-resolution radar image¹⁰ of the Earth-crossing asteroid 4179 Toutatis is shown in a. Is the interior of Toutatis a monolith or a multi-component structure, such as the fragmented aggregate shown in (b), contact binary (c) or 'rubble pile' (d)? Housen and colleagues' model¹ of asteroid 253 Mathilde suggests that some asteroids may be 'super' rubble piles — that is, they are so highly porous that craters form by compaction without ejection of any material.

are quite porous, either because their internal structure is fragmented or because they are loosely bound piles of material. Many asteroids have giant craters that are half their total size or larger. Craters of that size cannot form in solid rock without shattering the parent body to pieces², so gravity — even though it's weak — must be the dominant factor holding asteroids together. Most asteroids are now believed to be multi-component structures, as illustrated in Fig. 1. But with an internal strength so low that one might ignore it, and gravity so low one might ignore that too, one quickly runs out of physics to explain asteroid cratering. The key to the latest asteroid models is to consider both gravity and strength, however weak.

High porosity may turn out to be an asteroid's salvation: computer models and laboratory experiments have shown^{3,4} that because a fractured asteroid transmits impact energy so poorly, it is much more

difficult to break up than a monolith. Like the most flexible trees in a storm, composite asteroids are survivors. Further evidence that such 'rubble piles' are common is obtained from the low number of rapidly spinning asteroids. No known asteroid larger than a few hundred metres rotates faster than self-gravity can sustain⁵ — one would expect faster rotation if internal strength held them together.

Asteroid Mathilde, encountered in 1997 by the NEAR (Near Earth Asteroid Rendezvous) spacecraft (Box 1, overleaf), added a new twist to the puzzle. Mathilde is about 53 km in diameter and is the first outer-belt asteroid ever imaged⁶ (these are thought to be the most primitive asteroids): it is darker than charcoal, with a red hue suggesting a rich organic chemistry, but no evidence of water⁷. Its density was measured⁸ to be only 1.3 g cm^{-3} , which implies a porosity greater than 50% if it has the same composition as

Box 1: Asteroid interiors—the final frontier

Half a dozen asteroids have had close encounters with spacecraft. Gaspra, Ida, and Dactyl (Ida's moon) were seen by Galileo en route to Jupiter; Braille was glimpsed during the Deep Space 1 flyby; and Mathilde and near-Earth asteroid Eros were imaged by NEAR, which is returning for a much closer look at Eros during a year-long rendezvous beginning next February. Although NEAR will give us high resolution, multispectral images of the surface of Eros, and a wealth of other data, more direct sensing is needed

to settle the debate about what is inside. If the Deep Impact mission to Temple 1 gets the go-ahead, in 2005 we shall witness an energetic cratering event when its 500-kg copper payload slams into the comet's nucleus (perhaps not that different from a primitive asteroid) at 10.5 km s^{-1} . This is the cheapest way to look inside a comet: Deep Impact will investigate near-surface composition and structure, and may produce surprises if Housen *et al.*¹ are correct about compaction cratering. Another idea for exploring the

deep interiors of asteroids would be a double-payload mission, where one first images the target asteroid while distributing accelerometers upon the surface (first flyby), and then records data from the accelerometers as the asteroid responds to a projectile (second flyby). This would yield an acoustic picture of the asteroid's interior, in the same way that seismologists learned about the interior of the Earth earlier this century by listening to its tell-tale vibrations in response to earthquakes. E. A.

the most primitive meteorites (carbonaceous chondrites) found on Earth. As surprising as Mathilde's low density are the five or six giant, rimless craters (the largest is 33 km wide) that mark its surface: it looks as though some deity has taken large bites from a cosmic potato (see Fig. 1 on page 156). The lack of rims, or any noticeable crater debris, presents a paradox. Perhaps our assumptions are naive, but standard modelling⁹ predicts that such impacts will result in asteroid break-up rather than cratering (if gravity is negligible), or else that copious debris ejected from these vast craters will be deposited all around the asteroid (if strength is negligible). On Mathilde there is evidence for neither.

What might the explanation be? On seeing the first NEAR images of Mathilde and learning of its low density, the late Gene Shoemaker proposed that porosity might protect pre-existing craters from the violence of subsequent impacts. This explains how all these craters could survive, but it does not account for the lack of debris. Our group⁴ ran supercomputer simulations of asteroid impacts showing that the confinement of impact energy in a porous target could lead to ejection velocities greater than escape velocity, so that most crater material would be blasted from the asteroid leaving a clean surface behind.

Housen *et al.*¹ now put forward an exciting, if peculiar, alternative hypothesis. It also assumes that Mathilde has a rubble-pile structure, but invokes a vastly different cratering mechanism and suggests that highly porous asteroids might accumulate rather than lose mass during collisions. It is based on experimental data and scaling arguments. (Scaling, also known as dimensional analysis, is familiar to fans of Godzilla: dress a six-foot man as a 600-foot monster, and then slow down the action to a tenth of its original

speed to achieve realism.) In this case Housen *et al.* simulate large-scale impacts on low-gravity Mathilde by studying small-scale impacts in highly porous targets and using a centrifuge to achieve conditions of increased gravity. They equate 6-cm craters formed at 500 G (Earth's gravity is 1 G) to 33-km craters forming at 0.0009 G, and show that Mathilde's giant craters could be formed by compression of a highly porous target, and not as a result of ejected material. High porosity might be a transient feature of such asteroids, which become more dense with each impact.

To some, this will seem a wild idea. To others, especially those who remember how far-fetched the idea of rubble-pile asteroids

sounded ten years ago, it is certainly plausible. The truth may be a combination of factors, and has considerable implications. Not only can asteroid structure help to reveal how planets accumulated in the early Solar System, it must be understood before we can reliably divert an asteroid that might become a threat to Earth. If highly porous asteroids do crater by compaction, as Housen *et al.* propose, it might allow primitive asteroids to scoop up material like a celestial Pac-Man. That would make planetary accretion easier to understand, and is good news for modellers who want their planets to start to grow. But if, instead of being diverted or fragmented by a missile, an asteroid can simply absorb explosive energy, the conclusion is potentially bad news for planetary protection. Even so, we would have to live ten million years or more, on average, to experience the type of asteroid collision that could wipe out mankind. I am not betting on that horse. ■

Erik Asphaug is in the Department of Earth Sciences, Earth and Marine Sciences Building, University of California at Santa Cruz, California 95064, USA.

e-mail: asphaug@earthsci.ucsc.edu

1. Housen, K. R., Holsapple, K. A. & Voss, M. E. *Nature* **402**, 155–156 (1999).
2. Asphaug, E. & Melosh, H. J. *Icarus* **101**, 144–164 (1993).
3. Love, S. G., Hoerz, F. & Brownlee, D. E. *Icarus* **105**, 216–224 (1993).
4. Asphaug, E., Ostro, S. J., Hudson, R. S., Scheeres, D. J. & Benz, W. *Nature* **393**, 437–440 (1998).
5. Harris, A. W. *Lunar Planet Sci.* **27**, 493–494 (1996).
6. Veverka, J. *et al. Science* **278**, 2109–2112 (1997).
7. Rivkin, A. S., Clark, B. E., Britt, D. T. & Lebofsky, L. A. *Icarus* **127**, 255–257 (1997).
8. Yeomans, D. K. *et al. Science* **278**, 2106–2109 (1997).
9. Housen, K. R., Schmidt, B. R. & Holsapple, K. A. *J. Geophys. Res.* **88**, 2485–2499 (1983).
10. Ostro, S. J. *et al. Science* **270**, 80–83 (1999).

RNA interference

Policing rogue genes

Rueyling Lin and Leon Avery

Life is based on a social contract — genes work for the good of the organism, and they are reproduced. But certain rogue genes, called transposons, wantonly reproduce at the expense of the organism, inserting new copies of themselves all over the genome. Reporting in *Cell*, Tabara *et al.*¹ and Ketting *et al.*² now show that organisms have systems to hold transposons in check. They suggest that one of the clues that organisms use to detect illicit activity is double-stranded RNA, and their results could explain the reason for the mysterious phenomenon of RNA interference.

RNA interference was discovered by Guo and Kemphues³ as a method of inactivating genes in the germ line (the cell lineage that gives rise to offspring) of the nematode

worm *Caenorhabditis elegans*. In many systems, introduction of 'antisense' RNA — that is, RNA complementary to the 'sense' or messenger RNA (mRNA) — interferes with gene function. The antisense RNA was thought to hybridize to the appropriate mRNA transcript and block its translation to protein. But Guo and Kemphues noticed that if they injected sense RNA into *C. elegans* gonads, it inactivated the targeted gene as efficiently as the antisense RNA, ruling out hybridization to mRNA as the inactivation mechanism.

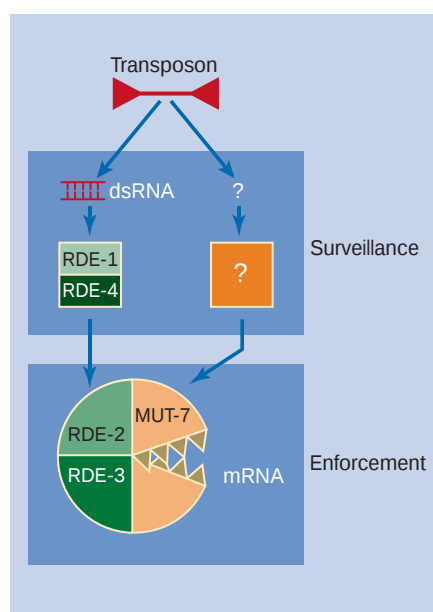
Fire *et al.*⁴ subsequently discovered that doubled-stranded RNA (dsRNA) works much better than either sense or antisense RNA alone. But RNA sequences that are complementary to introns (non-coding

sequences, which are not present in mRNA) do not interfere with gene expression. This suggests that RNA interference occurs after introns are removed. The interference seems to be catalytic rather than stoichiometric — a few molecules of dsRNA per cell are enough — and, in the cases examined^{4,5}, there was no detectable mRNA or protein for the targeted gene. These results suggest a mechanism involving enzymatic removal of mRNA. Surprisingly, RNA interference can even cross cell boundaries: although dsRNA is typically injected into the adult gonad, soaking worms in dsRNA⁶ or feeding them dsRNA-expressing bacteria⁷ also works.

RNA interference is by far the quickest and easiest method to inactivate *C. elegans* genes, and it has been widely used despite the mystery about how it works. It has also been reported in other organisms^{8–12}, suggesting that it may be evolutionarily conserved. However, these reports are relatively few, suggesting that RNA interference in other organisms is not as robust.

Obviously, the natural function of RNA interference is not to allow experimenters to inactivate genes. What, then, might it be? To address this question, Tabara *et al.*¹ looked for mutants defective in RNA interference. Normal worms grown on a lawn of bacteria expressing dsRNA for the *pos-1* gene produce dead embryos. But mutants defective in RNA interference produced viable offspring, allowing the authors to identify four so-called *rde* genes. Mutants defective in either *rde-1* or *rde-4* showed no other defects, indicating that RNA interference is not required for normal life. The *rde-1* gene encodes a protein belonging to a large family of unknown function, one member of which — the *sting* gene in the fruit fly *Drosophila melanogaster* — is needed to silence the repetitive germline-specific *Stellate* locus¹³. Unlike *rde-1* and *rde-4* mutants, however, worms mutated in the *rde-2* and *rde-3* genes showed other defects. These included activation of repetitive transgenes in the germ line (where transgenes are normally silenced), and mobilization of transposons. The link between RNA interference and transposons is intriguing, because the phenomenon works especially well in the germ line — where transposons are normally inactive¹⁴.

Ketting *et al.*² independently discovered this link. Many natural *C. elegans* isolates have high transposon activity whereas others, including the usual laboratory strain Bristol N2, do not¹⁵. Ketting and colleagues isolated 30 mutations that activate transposons in the Bristol N2 germ line. Of these, they found that 22 also caused defects in RNA interference. One of the genes identified, *mut-7*, encodes a protein that is homologous to an RNA-degrading enzyme, RNaseD. This is consistent with the idea that RNA interference involves enzymatic degra-



dation of mRNA, and that mRNA degradation prevents transposon activity.

All of these data suggest that RNA interference exists to suppress transposon activity. Two further arguments support this model. First, it explains why RNA interference is most effective in the germ line: transposition in the germ line damages the DNA that is passed on to offspring, whereas transposition outside the germ line is less serious. Second, active transposons are expected to produce dsRNA. Transposons often insert into genes and are transcribed into RNA. Most transposons have inverted-repeat sequences at their ends that, when transcribed into RNA, will hybridize to form dsRNA. A transposon without inverted repeats can also produce dsRNA if it is transcribed in opposite directions (as a result, for instance, of inserting into several genes). In fact, RNA interference is only one of the many responses triggered by dsRNA in living organisms. Perhaps all of these responses evolved as a defence against rogue genes.

Intriguing though this explanation of RNA interference is, there is a problem — mutations in the *rde-1* and *rde-4* genes eliminate RNA interference but they do not activate transposons. The model can be reconciled with this awkward fact if we assume that dsRNA is not the only clue used to detect transposons. Suppose the genome police have 'surveillance' and 'enforcement' divisions (Fig. 1). The RDE-1 and RDE-4 proteins could be part of the surveillance team, their function being to detect dsRNA. Other (currently unknown) molecules detect alternative clues of transposon activity. When alerted, the surveillance team informs the enforcement division (the RDE-2, RDE-3 and MUT-7 proteins), which shuts down the illicit activity by degrading transposon mRNA. The enforcement division is necessary to suppress transposon activity, but

those components of the surveillance team that detect dsRNA are not, because transposons can be detected via the alternative clues.

Whether or not this explanation is true, three things are clear. First, RNA interference is not necessary for normal life in *C. elegans*. Second, there is some link between RNA interference and transposon activity. Finally, genetic analysis of RNA interference will be straightforward, and more answers should be coming soon. But the systems uncovered could be complex. After all, worms have been fighting transposons for millions of years, and transposons have almost certainly evolved ways to evade detection — and the genome police new ways of detecting them. ■

Rueyling Lin and Leon Avery are in the Department of Molecular Biology, University of Texas Southwestern Medical Center, 6000 Harry Hines Boulevard, Dallas, Texas 75390-9148, USA.
e-mails: leon@eatworms.swmed.edu
rlin@hamon.swmed.edu

1. Tabara, H. *et al.* *Cell* **99**, 123–132 (1999).
2. Ketting, R., Haverkamp, T., van Luenen, H. & Plasterk, R. *Cell* **99**, 133–141 (1999).
3. Guo, S. & Kemphues, K. J. *Cell* **81**, 611–620 (1995).
4. Fire, A. *et al.* *Nature* **391**, 806–811 (1998).
5. Montgomery, M. K., Xu, S. & Fire, A. *Proc. Natl Acad. Sci. USA* **95**, 15502–15507 (1998).
6. Tabara, H., Grishok, A. & Mello, C. *Science* **282**, 430–431 (1998).
7. Timmons, L. & Fire, A. *Nature* **395**, 854 (1998).
8. Waterhouse, P. M., Graham, M. W. & Wang, M. B. *Proc. Natl Acad. Sci. USA* **95**, 13959–13964 (1998).
9. Ngo, H., Tschudi, C., Gull, K. & Ullu, E. *Proc. Natl Acad. Sci. USA* **95**, 14687–14692 (1998).
10. Kennerdell, J. R. & Carthew, R. W. *Cell* **95**, 1017–1026 (1998).
11. Lohmann, J. U., Endl, I. & Bosch, T. C. *Dev. Biol.* **214**, 211–214 (1999).
12. Sanchez-Alvarado, A. & Newmark, P. *Proc. Natl Acad. Sci. USA* **96**, 5049–5054 (1999).
13. Schmidt, A. *et al.* *Genetics* **151**, 749–760 (1999).
14. Plasterk, R. & van Luenen, H. in *C. elegans II* (eds Riddle, D., Blumenthal, T., Meyer, B. & Priess, J.) 97–116 (Cold Spring Harbor Laboratory Press, 1997).
15. Moerman, D. G. & Waterston, R. H. *Genetics* **108**, 859–877 (1984).



100 YEARS AGO

On the shore of the island of Hadod, latitude $68^{\circ} 40'$ about, in Vesterdaalen, north of Lofoten, there was found, probably in the autumn of 1897, a wooden ball, $4\frac{1}{2}$ centimetres in diameter, covered by a thin layer of gum. The ball is of fine workmanship, and just able to float in the water. Circles are engraved upon four parts, and form small rhumbs over the whole surface; and on two places there is engraved with Latin Majuscles the name *Melfort*. Perhaps some of your readers can say from whence this ball has come. I am writing to the man who has the ball now, to ask him to send it to me.

From *Nature* 9 November 1899.

50 YEARS AGO

The speed of the meteors is so great that as they rush into the atmosphere of the earth they burn away at heights of about sixty miles. Until recently, the streaks of light produced by the burning meteors provided the only method by which astronomers could study this phenomenon. Such observations are hindered or prevented by cloud and moonlight, and are impossible in daylight. On the other hand, the burning meteor leaves behind it a dense trail of electrons which can reflect radio waves ... An important consequence of the radio observations has been the discovery of great meteoric activity in summer day-time. The well-known visual meteor showers ... usually last for a few nights. The summer daytime showers detected by the radio method are far more extensive and attain higher rates. The daylight activity, which in May comes from the direction of Pisces, develops rapidly, extending over a wide belt stretching towards the sun, and continues as a great succession of meteor showers until late August. The origin of the streams of meteoric debris is a most important problem of meteor astronomy. It is almost certain that most of them move in orbits around the sun. A few are associated with comets; but it is unlikely that this is the case with all of them. There is difference of opinion as to whether the meteoric debris is all localized in the solar system or whether some comes from interstellar space. The radio observations now being made will help answer these questions. The accurate determination of speeds will resolve the problem as to whether any meteors originate from outside the solar system.

From *Nature* 12 November 1949.

Genetics

Distorting sex ratios

Keith R. Willison

During sperm production in mammals, the process of meiosis generates spermatozoa that are genetically different but functionally equivalent. So, with respect to the sex chromosomes, for example, equal numbers of sperm are produced carrying either an X or a Y chromosome. Because X and Y sperm are equally capable of fertilizing an egg, equal numbers of male and female embryos are produced, and the mammalian sex ratio at birth is 50%.

All other pairs of chromosomes show this 50% transmission ratio, but there are rare exceptions. For example, some mice contain naturally occurring variants of chromosome 17 — the so-called *t*-haplotypes — that do not obey this rule. Remarkably, sperm carrying these unusual chromosomes can propagate themselves with transmission ratios as great as 99% in their own favour. It was not known how this process, known as transmission ratio distortion (TRD), is achieved at a molecular level. But on page 141 of this issue, Herrmann and colleagues¹ describe a unique protein kinase encoded by the *t*-haplotype. This kinase affects a signal-transduction cascade that probably controls the speed and directionality of sperm as they make their long journey up the female reproductive tract.

Within a mouse *t*-haplotype, several hundred functionally unrelated genes have been locked together by a series of chromosomal inversions that prevent recombination with wild-type chromosomes². All *t*-haplotypes derive from a single common ancestor, and wild mice carrying these chromosomes can be found all over the world. The *t*-haplotype variants (known as t^{w12} , t^{w32} and so on) differ in the functionally unrelated genes, but not in the TRD system. Presumably, it is the common TRD system that allowed the successful propagation of *t*-haplotypes.

Over 40 years ago, Mary Lyon began a genetic study of TRD in the t^6 haplotype of laboratory mice. She eventually developed a model for TRD, involving interactions between three or more genetic loci^{3–5}. According to this model, several loci — the *t*-complex distorters, *Tcd* — act on the single *t*-complex responder locus, *Tcr*, to distort transmission ratios (Fig. 1a). Normally, of course, all the component genes of the TRD system are encoded in the *t*-region of chromosome 17, and are locked together by the chromosomal inversions. But experiments have shown that the distorters (*Tcds*) can act in *trans* on the *cis*-acting *Tcr* responder³. Put more simply, this means that the

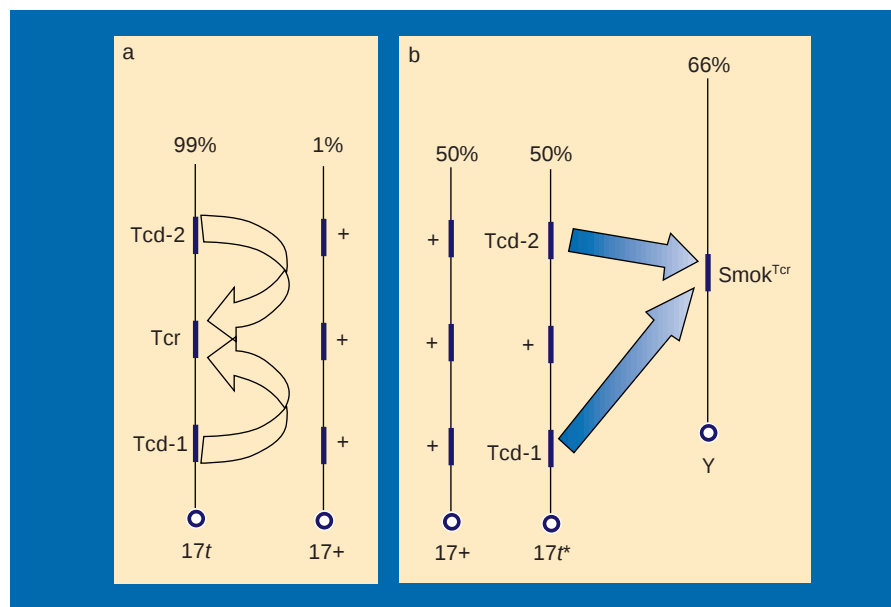


Figure 1 Genetic models of transmission ratio distortion. **a**, Naturally occurring configuration of genetic elements required for 99% transmission of a *t*-haplotype chromosome (17*t*) compared with a wild-type chromosome (17+). Two or more distorters (*Tcd*-1 and *Tcd*-2) act on the responder (*Tcr*) (simplified model from refs 3,4). **b**, Experimental configuration of genetic elements created by Herrmann and colleagues¹, which produce twice as many male mice as females. The natural distorters are carried by a modified *t*-haplotype lacking *Tcr* (17*t**), and they act in *trans* on a transgenic *Tcr* DNA construct carried on the mouse Y chromosome.

Box 1 : Sperm sexing and selection

There are obvious advantages to being able to pre-select the gender of farm animals — in breeding cattle for milk production, for example¹⁰. Current techniques use sperm sorting followed by artificial insemination, based on the difference in DNA content between the X and Y chromosomes. In mammals, sperm bearing X chromosomes have 2.8–7.5% more DNA than

those with Y chromosomes. Although this process works (it can produce sex ratios of 85–90% in either direction), it is inefficient and not generally applicable. Assuming the Smok-kinase system is evolutionarily conserved in mammalian sperm production, it may be possible to use the *Tcr* responder to manipulate the transmission ratio of sex chromosomes

genetically. Male cattle carrying *Tcr* on the Y chromosome might be expected to transmit the X chromosome preferentially. The *Tcr* responder could also be useful in mouse genetics. In genetic crosses where many construct genes are being segregated, tagging constructs or chromosomes with *Tcr* cassettes could give better segregation in subsequent manipulations.

K. R. W.

chromosome carrying *Tcr* is transmitted at a high ratio. It is the identification¹ of this *cis*-acting element, *Tcr*, that is now so exciting.

The location of *Tcr* had previously been restricted to a 155-kilobase region of the *t*-haplotype⁶. Within this region, Herrmann and colleagues¹ found a protein kinase gene spanning 80 kilobases. This *Tcr* kinase gene is a fusion of two kinases — a ribosome S6 kinase 3 (*Rsk3*), and a hitherto unknown kinase related to the microtubule-associated protein (MAP)/microtubule affinity-regulating kinase (MARK) family⁷. This family of serine/threonine kinases is known to phosphorylate MAPs and cause them to detach from microtubules. Several members of the new, MARK-related family are found on the *t*-haplotype chromosome, and also on wild-type chromosome 17. They have been designated 'sperm-motility kinases' (gene symbol *Smok*).

The *Tcr* gene probably derived initially from a rearrangement between a *Smok* gene and the neighbouring *Rsk3* gene. Presumably, during evolution of the *t*-haplotype, mutation and selection generated the dominant genetic factor, *Smok^{Tcr}*, or *Tcr* for short. Although there are many *Smok* genes in this region of chromosome 17, there is only one *Tcr* kinase. Because the generation and subsequent selection of *Tcr* seems so unlikely, it is not surprising that TRD only evolved once, and that all *t*-haplotypes are so similar.

To show that the *Tcr* kinase corresponds to the genetic element defined by Lyon, Herrmann *et al.*¹ constructed *Tcr* transgenic mice in which *Tcr* has randomly incorporated into various mouse chromosomes. The authors found that *Tcr* behaves exactly as predicted by Lyon's model^{3,4}, giving high transmission of the *Tcr*-containing chromosome in the presence of *Tcds*, yet less than 50% transmission in their absence. One *Tcr* transgenic strain has *Tcr* incorporated on the Y chromosome, and this produces twice as many males as females (Fig. 1b and Box 1).

But what do these results mean physio-

logically? The *t*-haplotypes are known to compromise the flagella that power sperm swimming⁸, rather than sperm production *per se*. The flagellum, found in the sperm tail, is composed mainly of microtubules (made from a protein called tubulin) and other components that regulate their activity, such as dyneins. Dyneins allow special paired microtubules to slide against each other, then cross-linking proteins help to convert this sliding to a bending motion that allows the sperm tail to beat.

Herrmann *et al.*¹ speculate that the *Tcr* system targets the dyneins. During sperm formation, cells containing a normal copy of chromosome 17 are thought to be damaged by distorter proteins produced by cells that contain the *t*-haplotype chromosome 17.

Condensed matter

Through a glass, lightly

A. Lindsay Greer

Glass formation arises when a liquid is cooled below its freezing point without crystallization. It has been shown that the glass transition is kinetic rather than thermodynamic in origin, but how atomic motion becomes stifled below the glass transition temperature (T_g) is a long-standing problem. On page 160 of this issue¹, Tang *et al.* demonstrate with particular clarity that atomic motion below T_g involves single-atom hopping, whereas motion above T_g is more collective. The idea of such changes in mechanism around T_g is far from new, but this paper helps remove speculation about the exact mechanism. Tang and colleagues' contribution may be compared with the first view of Earth from space — not surprising to see that it is round, but very nice to have the picture.

The new work concerns the diffusion of atoms in bulk metallic glass, but has implications for glass-forming systems in general.

This is possible because developing sperm grow together in a syncytium, in which groups of cells remain connected to each other through large junctions, allowing macromolecules to diffuse freely⁹. The action of *Tcr*, on the other hand, could be to protect those sperm cells that contain the *t*-haplotype from the harmful action of its own distorters — that is, *t*-haplotypes encode both the poison and the antidote. But it is not clear how the antidote might function in only half the cells in the syncytium, even with the knowledge that *Tcr* is a protein kinase and that it is expressed in the right cells at the right time. It is to be hoped that cell-biological analysis of the system will tell us how *Tcr* can act in *cis*, and how it interacts with the newly discovered kinase cascade that regulates the sperm flagellar apparatus.

Keith R. Willison is at the Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London SW3 6JB, UK.
e-mail: willison@icr.ac.uk

- Herrmann, B. G., Koschorz, B., Wetz, K., McLaughlin, K. J. & Kispert, A. *Nature* **402**, 141–146 (1999).
- Silver, L. M. *Trends Genet.* **9**, 250–254 (1993).
- Lyon, M. F. *Cell* **37**, 621–628 (1984).
- Lyon, M. F. *Cell* **44**, 357–363 (1986).
- Willison, K. R. & Lyon, M. F. *Int. J. Dev. Biol.* (in the press).
- Bullard, D. C., Ticknor, C. & Schimenti, J. C. *Mamm. Genome* **3**, 579–587 (1992).
- Drewes, G., Ebner, A., Preuss, U., Mandelkow, E. M. & Mandelkow, E. *Cell* **89**, 297–308 (1997).
- Olds-Clarke, P. & Johnson, L. *Dev. Biol.* **155**, 14–25 (1993).
- Braun, R. E., Behringer, R. R., Peschon, J. J., Brinster, R. L. & Palmiter, R. D. *Nature* **337**, 373–376 (1989).
- Johnson, L. A. *Deutsch Tierarztl. Wochenschr.* **103**, 288–291 (1996).

Although metallic glasses have been around for some 40 years², they are still parvenus. But with the discovery of more stable examples (that is, with greater resistance to crystallization) a broader range of experimental studies is now possible. The more stable metallic glasses are alloys, such as $Zr_{41.2}Ti_{13.8}Cu_{12.5}Ni_{10}Be_{22.5}$ (the composition is in atomic %) studied by Tang *et al.*¹; they permit glass formation at a slower cooling rate and therefore in greater bulk. When heated through the glass transition, the new metallic glasses yield a supercooled liquid state of unusual stability for metals, permitting reliable measurements of atomic motion in this state for the first time.

The essence of the new study is the comparison of hopping rates for beryllium atoms (Be), as measured by nuclear magnetic resonance (NMR), with diffusive transport rates of Be measured by composition profiling. The problem with NMR measurements in

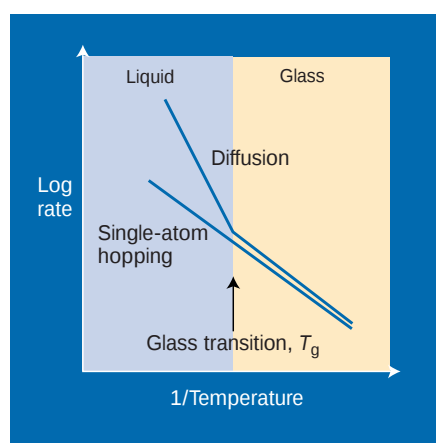


Figure 1 Temperature dependence of the rates of diffusion and single-atom hopping for Be atoms in a bulk metallic glass studied by Tang *et al.*¹. Below T_g the rate of single-atom hopping and diffusion show the same temperature dependence, suggesting that atomic jumps are responsible for long-range diffusion in the glassy state. Above T_g single-atom hopping cannot account for all the observed diffusion, and a collective motion of more than one species of atom is required to explain long-range diffusion in the supercooled liquid state.

metals is that the timescale of atomic motion near the glass transition is very long compared with the relevant nuclear spin-lattice relaxation time. In the case of the ^9Be atoms studied by Tang *et al.* the relaxation time is long enough to allow detection of atomic jump rates. The authors use an unusual spin alignment echo (SAE) technique to detect Be hopping in the bulk metallic glass. They consider with care the nature of the echo decay from the ^9Be atoms; its pure exponential form confirms that the SAE technique detects unconstrained hopping that contributes fully to long-range atomic transport and is not merely a rattling within local atomic cages. Although the arguments are subtle, they also make a clear case that SAE detects hopping of single Be atoms and not the collective motion of a cluster of atoms.

The results are compared in Fig. 1. Below T_g ($\sim 620\text{ K}$ in this case), the rates of single-atom hopping and diffusion show the same temperature dependence. Importantly, the single-atom hopping is at exactly the right rate to explain the observed diffusion. So below T_g the authors conclude that diffusion is mainly by single-atom hopping. Above T_g the diffusion increases much more rapidly with temperature. Such a break in behaviour is well known³, and is exhibited by a range of diffusing species. It had been suggested in earlier related work⁴, that this must reflect increased hopping rates above T_g , perhaps enhanced by evolving atomic configurations around the hop site. The new work shows definitively that this is not the case — the rate of hopping shows a continuous temperature dependence with no break at T_g (Fig. 1).

The hopping-rate data now make it possible to move beyond speculation about diffusion mechanisms. Together, the SAE and diffusion studies indicate that the motion of Be atoms above T_g is due to two diffusion processes, one of which is single-atom hopping. The other process (which dominates as soon as the temperature is raised above T_g) must be collective, because the SAE technique is not sensitive to collective motion. It now appears that diffusion above T_g is mainly by collective motion of a small number of nearest neighbour atoms, rather than by single-atom hopping.

The collective motion of atoms above T_g (for which evidence also comes from comparing the rates of diffusion of different isotopes⁵) may lead to the motions of all species being coupled. As suggested recently by Johnson⁶, the change in diffusive behaviour at T_g can then be described as a decoupling in which the diffusion below T_g becomes faster than would be predicted from the viscosity of the glass. The break is less evident for larger, slower species. The Be atoms featured in the new work are an intriguing intermediate case. They are small enough to tread lightly through the glass below T_g , yet their diffusion is strongly coupled to the other species above T_g . Were a still smaller species to be studied, it is likely that its diffusion would be dominated by single-atom hopping throughout, with a temperature dependence unbroken at T_g . For example, just such

behaviour is shown by water diffusing in large-molecule glasses (carbohydrates and polymers)⁷. It seems that Be in the bulk metallic glasses can be regarded as a full member of the dense random-packed structure, and not as an interstitial species (like water).

One view of the glass transition is that the same transport processes might operate above and below T_g , with only the 'freezing out' of structural changes giving different temperature dependence below T_g . This view is defeated by Tang and colleagues' data, which show a clear change in diffusion mechanism at T_g . The application of NMR to diffusion studies in other glasses is well established⁸. The study of metallic glasses promises further advances in exploring the range of coupling and decoupling behaviour that makes up the glass transition. No more seeing through a glass, darkly? ■

A. Lindsay Greer is in the Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK.

e-mail: alg13@cam.ac.uk

1. Tang, X.-P., Geyer, U., Busch, R., Johnson, W. L. & Wu, Y. *Nature* **402**, 160–162 (1999).
2. Klement, W. *et al.* *Nature* **187**, 869–870 (1960).
3. Wenner, F. *et al.* *Defect Diff. Forum* **143**, 831–835 (1997).
4. Geyer, U. *et al.* *Phys. Rev. Lett.* **75**, 2364–2367 (1995).
5. Ehmlich, H. *et al.* *Phys. Rev. Lett.* **80**, 4919–4922 (1998).
6. Johnson, W. L. *MRS Bull.* **24**, 42–56 (1999).
7. Aldous, B. J. *et al.* *J. Mater. Sci.* **32**, 301–308 (1997).
8. Oksanen, C. A. & Zografi, G. *Pharm. Res.* **10**, 791–799 (1993).

Cell cycle

Dual control of mitotic exit

Susanne Prinz and Angelika Amon

The cell cycle is a precisely timed sequence of events known as S phase, G2, mitosis and G1. During S phase, the genetic material is duplicated. Then, after a rest phase (G2), it is segregated equally between two daughter cells (mitosis). Cells finally enter another resting phase, G1, before the cycle starts over again.

The transition to G1 is known as exit from mitosis. It is regulated by a protein called Cdc14, which acts as a phosphatase (that is, it dephosphorylates specific protein targets)¹. Another protein, Cdc20, is also known to promote exit from mitosis^{2,3}, but it is not clear how. On page 203 of this issue, Shirayama *et al.*⁴ show that Cdc20 has a dual function in triggering this final cell-cycle transition. By facilitating the degradation of a protein called Pds1, Cdc20 releases Cdc14 from its prison in the nucleolus. Then, by inducing degradation of Clb5 (an antagonist of Cdc14), Cdc20 ensures that Cdc14 can efficiently dephosphorylate its targets.

Key cell-cycle transitions — be they

entry into S phase or progression through mitosis — are mediated by specialized protein-degradation systems. These consist of the so-called ubiquitination machinery, which attaches a peptide (ubiquitin) onto the protein to be degraded, and a protease, the 26S proteasome, that degrades this ubiquitin-tagged protein⁵. Mitosis is regulated by the anaphase-promoting complex (APC) ubiquitination machinery, along with a family of APC activators, Cdc20 and Cdh1/Hct1 (ref. 6). An APC–Cdc20 complex initiates the chromosome-separation phase (known as 'anaphase') by ubiquitinating the anaphase inhibitor Pds1 (Fig. 1, overleaf). But Cdc20 is also involved in mitotic exit^{2,3}, as cells that lack *CDC20* and *PDS1* go through anaphase but cannot exit mitosis and enter G1.

Exit from mitosis is ultimately triggered by inactivation of mitotic cyclin-dependent protein kinases (CDKs). These consist of a catalytic protein-kinase subunit and a regulatory cyclin subunit. The APC, together

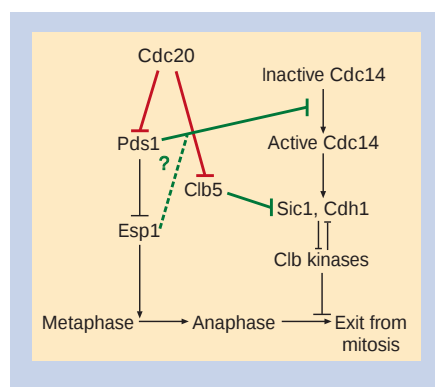


Figure 1 Two-step regulation of exit from mitosis by Cdc20, as proposed by Shirayama *et al.*⁴. After completion of metaphase (the stage of mitosis when chromosomes condense and a mitotic spindle forms), Cdc20, together with the anaphase-promoting complex, ubiquitinates Pds1. Degradation of Pds1 frees its partner Esp1, which then initiates the partitioning of the genetic material between daughter cells. Degradation of Pds1 also allows the release of Cdc14 from the nucleolus. Whether Esp1 mediates the Pds1-dependent inhibition of Cdc14 is not clear (dotted line). Cdc20 also degrades the Cdc14 antagonist Clb5. The Clb5–Cdc28 kinase phosphorylates Sic1 and Cdh1 more rapidly than Cdc14 dephosphorylates them. So, Cdc20 enables Cdc14 to dephosphorylate its targets Sic1 and Cdh1, which, in turn, allows these proteins to become active and destroy mitotic CDKs, resulting in exit from mitosis.

with its activator Cdh1, induces degradation of the mitotic cyclins Clb1, Clb2, Clb3 and Clb4, and, as a result, inactivates the mitotic CDKs. Binding of Sic1 to the kinase complex also downregulates mitotic CDKs and triggers exit from mitosis. So Cdc14 works by dephosphorylating Cdh1 and Sic1, facilitating both Clb degradation and Sic1 accumulation¹. But how is Cdc14 inactivated when it is not needed during G1, S phase and metaphase? It is then sequestered in the nucleolus by Cfi1/Net1 (ref. 1). Only during anaphase, when a signal-transduction pathway termed the mitotic-exit pathway^{7,8} releases it from the nucleolus, can Cdc14 reach and dephosphorylate its targets.

To determine how Cdc20 regulates exit from mitosis, Shirayama *et al.*⁴ did a genetic selection based on the premise that exit from mitosis requires degradation of a certain factor by Cdc20, and that the absence of this factor will allow cells lacking *CDC20* and *PDS1* to exit mitosis, divide and form colonies. Using this mutant hunt the authors identified a gene called *CLB5*, which encodes a cyclin required for progression through S phase. The Clb5 cyclin, when complexed with the protein kinase Cdc28, is important for promoting S phase⁷.

Having confirmed that Clb5 is indeed degraded by the APC–Cdc20 complex, Shi-

rayama and colleagues worked out how the degradation of Clb5 and Pds1 regulates exit from mitosis. They examined the levels of mitotic CDKs and localization of Cdc14 in *cdc20 clb5* and *cdc20 pds1* double mutants, and in *cdc20 clb5 pds1* triple mutants. As expected, neither *cdc20 pds1* nor *cdc20 clb5* double mutants could inactivate mitotic CDKs (and, therefore, exit mitosis), whereas the *cdc20 clb5 pds1* triple mutant could. Interestingly, though, the two double mutants differed with respect to Cdc14 localization. Whereas Cdc14 remained in the nucleolus in *cdc20 pds1* mutants, the protein was released in *cdc20 clb5* mutants. The authors concluded that Cdc20 is required at two steps during the initiation of mitotic exit. In the first, Cdc20-mediated degradation of Pds1 releases Cdc14 from the nucleolus. And in the second, Cdc20 degrades Clb5, allowing Cdc14 to dephosphorylate Cdh1 and Sic1 after it has been released from the nucleolus.

How does degradation of Pds1 facilitate liberation of Cdc14 from the nucleolus? And how does destruction of Clb5 enable Cdc14 to dephosphorylate critical cell-cycle targets? Pds1 is known to inhibit the onset of anaphase by inhibiting another protein called Esp1 (Fig. 1)⁸. So Shirayama *et al.* determined whether Pds1 prevented Cdc14 release by inhibiting Esp1. They found that *esp1* mutants showed, at best, a modest delay in the inactivation of mitotic CDKs. This suggests that Pds1 inhibits Cdc14 release by means other than inhibition of Esp1. But other reports point to the involvement of Esp1 in inactivating mitotic CDKs⁹. Irrespective of whether Pds1 inhibits Cdc14 release by blocking Esp1 or some other, unidentified protein, it will be interesting to see if, and how, Pds1 interferes with the mitotic-exit pathway¹.

To explain how Clb5 inhibits exit from mitosis, Shirayama *et al.* propose that the

Clb5–Cdc28 complex phosphorylates Cdh1 and Sic1 at a rate that exceeds Cdc14's dephosphorylation activity towards these proteins. Clb5 seems to be the only Clb cyclin that can antagonize Cdc14 effectively. For instance, when the authors eliminated other Clb cyclins, such as Clb1, Clb3 or Clb4, the *cdc20 pds1* double mutants could no longer exit mitosis. By working out why only Clb5 seems to inhibit Cdc14, we may learn not only how exit from mitosis is regulated, but also how Clb kinases differ in their ability to select and phosphorylate different substrates.

We now have a good understanding of how Cdc20 regulates exit from mitosis. The next challenge will be to determine how Cdc20's substrates — Pds1 and Clb5 — regulate the activity of Cdc14. Our ultimate goal will be to understand how cells ensure that exit from mitosis is not initiated before the chromosomes have segregated. A prime candidate for coupling these two events is Pds1, which inhibits both anaphase and release of Cdc14 from the nucleolus. But cells that lack *PDS1* still exit mitosis only after the chromosomes have separated completely⁴. So the pathways that regulate sister-chromatid separation in the absence of *PDS1*, and that couple this event to exit from mitosis, still need to be identified. ■

Susanne Prinz and Angelika Amon are at the Center for Cancer Research, Massachusetts Institute of Technology, 40 Ames Street, Cambridge, Massachusetts 02139, USA.
e-mail: angelika@mit.edu

1. Morgan, D. O. *Nature Cell Biol.* **1**, E47–E53 (1999).
2. Yamamoto, A. *et al.* *J. Cell Biol.* **133**, 99–110 (1996).
3. Shirayama, M. *et al.* *EMBO J.* **17**, 1336–1349 (1998).
4. Shirayama, M. *et al.* *Nature* **402**, 203–207 (1999).
5. Peters, J.-M. *Curr. Opin. Cell Biol.* **10**, 759–768 (1998).
6. Zachariae, W. *et al.* *Genes Dev.* **13**, 2039–2058 (1999).
7. Nasmyth, K. *Curr. Opin. Cell Biol.* **2**, 166–179 (1993).
8. Ciosk, R. *et al.* *Cell* **93**, 1067–1076 (1998).
9. Tinker-Kulberg, R. L. *et al.* *Genes Dev.* **13**, 1936–1949 (1999).

Plant genetics

Knocking out nodules

Herman P. Spaink

Plant genetics has had a great impact over the past 150 years. In the late nineteenth century Gregor Mendel discovered the basic principles of inheritance using the pea (a leguminous plant). Then, in the 1940s, Barbara McClintock's pioneering work on maize uncovered the existence of mobile DNA elements called transposons. In the 1970s, DNA transfer from *Agrobacterium tumefaciens*¹ allowed foreign genes to be efficiently introduced into the plant genome. And the development of simple, rapid methods to obtain transgenic offspring from

Arabidopsis thaliana plants has allowed large collections of mutants to be generated. But genetics has been difficult in other species of plant, for which there are no such rapid methods.

On page 191 of this issue, Schauser *et al.*² show that this lack of methodology will no longer be the main barrier for genetic studies of leguminous plants. The authors have shown that a transposon called *Ac*, which is derived from maize, can be successfully used to disrupt genes in the leguminous plant *Lotus japonicus*. This is good news for those

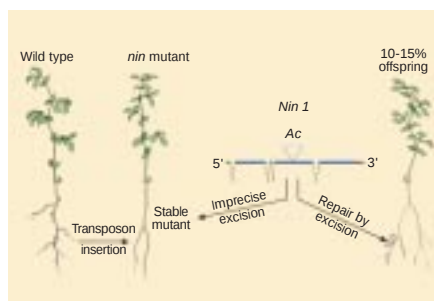


Figure 1 Strategy used by Schauser *et al.*² to characterize the 'nodule inception', *nin*, mutant. By inserting a mobile DNA element (transposon) called *Ac* into wild-type *Lotus japonicus*, the authors generated mutant plants with *Ac* inserted into the *nin* gene. This transposon can excise itself from the DNA. If it does so imprecisely, a mutation remains in the *nin* gene. But if it excises precisely, the gene can be repaired and the plants return to a wild-type phenotype. This is a powerful technique for generating mutations in plants.

studying root nodulation in the legume family. Root nodules are the result of a symbiotic (mutually beneficial) interaction between plants and various soil bacteria, collectively called rhizobia¹. Within these nodules, the bacteria fix atmospheric nitrogen into ammonia. This means that leguminous plants do not depend on fixed nitrogen in the soil — usually a major growth limitation.

Two decades of research into the molecular basis of this symbiosis have shown that it is due to an exchange of signal molecules. These molecules are recognized specifically by the host plants and their bacterial partners³, and they include the bacterially produced nodulation factors, which can trigger the roots of leguminous plants to form a root nodule⁴⁻⁶. The bacterial genes involved in the symbiosis have been studied in detail⁷, stimulating research into the genomes of rhizobia in general (for example, the sequence of the entire *Sinorhizobium meliloti* genome is close to completion⁸).

From the plant side, far less is known of the genes (and corresponding functions) involved in the symbiosis. This is mainly due to the technical bottlenecks in plant genetics. Various plants have been identified with an impaired capacity to be nodulated or infected by rhizobia⁹, but these mutants were derived by chemical mutagenesis. So the resulting defects (such as point mutations) are still difficult to identify.

The transposition approach used by Schauser *et al.*² now makes it very easy to identify mutated genes, because they are tagged by the known nucleotide sequence of the transposon. The strategy yielded a mutant that could not form root nodules, so was called a *nin* (for 'nodule inception') mutant (Fig. 1). Interestingly, the mutant

does not seem to be affected in any other aspect of plant development — such as root, shoot, leaf, flower or seed development — as long as nitrogen nutrients are supplied externally. This suggests that the mutated gene is dedicated to root-nodule formation.

The authors went on to show that the genetic defect in the *nin* mutants is indeed caused by integration of the transposon. To do this, they used the fact that the *Ac* transposon can excise spontaneously from the affected gene at a certain frequency. They found that when the transposon was precisely excised, mutant plants reverted to the wild-type phenotype (Fig. 1). Then, because the transposon acted as a tag for the mutated gene, Schauser *et al.* could easily obtain the nucleotide sequence of the complete gene using established procedures. This sequence showed that the *Nin* protein is probably a transcription factor, which may have a membrane-spanning domain. The *Nin* gene seems to be related to the *mid* ('minus dominance') genes from the alga *Chlamydomonas*. These genes regulate sexual reproduction in response to short supplies of nitrogen. Because root nodulation is also regulated by nitrogen limitation, this relationship might point to a conserved function for this type of regulator in evolution.

The identification of this hitherto-unknown regulator of nodule inception is a landmark discovery that will influence future studies in the legume field. This kind of approach will be strengthened by linking it to detailed analyses of leguminous-plant genomes. Two international consortia are currently obtaining nucleotide-sequence data from *L. japonicus* and another model legume, *Medicago truncatula*. Such large-scale projects are important owing to the role of leguminous plants in studies on the interaction of plants with symbiotic mycorrhizae⁸ (which are involved in the supply of phosphorus), or in research on the various plant pathogens that are specific for the leguminous family. It will be exciting to follow research in this area, and to see the vast amount of biochemical and physiological data obtained as leguminous plants are linked to plant genetics. ■

Herman P. Spaiink is at the Institute of Molecular Plant Sciences, Leiden University, 2333 AL Leiden, The Netherlands.

e-mail: spaiink@rubiim.leidenuniv.nl

1. Spaiink, H. P., Kondorosi, A. & Hooykaas, P. J. I. (eds) *The Rhizobiaceae* (Kluwer, Dordrecht, 1998).
2. Schauser, L., Roussis, A., Stiller, J. & Stougaard, J. *Nature* **402**, 191–195 (1999).
3. Fisher, R. F. & Long, S. R. *Nature* **357**, 655–660 (1992).
4. Lerouge, P. *et al.* *Nature* **344**, 781–784 (1990).
5. Truchet, G. *et al.* *Nature* **351**, 670–673 (1991).
6. Spaiink, H. P. *et al.* *Nature* **354**, 125–130 (1991).
7. Freiberg, C. *et al.* *Nature* **387**, 394–401 (1997).
8. Keen, N., Gelvin, S. & Long, S. *Mol. Plant–Microbe Int.* **12**, 835–838 (1999).
9. Penmetta, R. V. & Cook, D. R. *Science* **275**, 527–530 (1997).

Daedalus

Asteroidal defence

One day, a stray asteroid or comet may smash into the Earth, when we shall go the way of the dinosaurs. Pessimists advocate a close watch on asteroidal neighbours, and the building of a vast nuclear missile to disrupt the incomer before it hits the Earth. Daedalus has a neater scheme.

He wants to put an asteroidal 'anti-missile' in a high Earth orbit. If a potential invader were sighted, we would arrange for the guard asteroid to collide with it. An asteroid already in orbit would patrol near-Earth space very densely. Even a slight nudge, if well calculated, could guarantee a successful interception, shattering both asteroids.

But how to capture a guard asteroid in the first place? Many asteroids, comets and planetesimals contain hydrogen, water, ammonia or methane ices, which can deliver useful thrust by direct evaporation. Indeed, it has been suggested that an asteroid could be steered by flying a big concave mirror alongside it, to focus sunlight on the icy surface. The resulting plume of hot gas would deflect its orbit. Daedalus goes even further. Most asteroids, he claims, are not loose rubble piles of the sort discussed by E. Asphaug on page 127. They have a compressed centre, in which the ices are converted to energetic hydroxonium, ammonium or methylammonium 'alloys'. Release the over-pressure, and these compounds will decompose to hot pressurized gas. A space probe that landed on such an asteroid and drilled down to its energetic core would release a gas-jet that would turn the asteroid into a natural rocket. With crafty enough steering, it could be detached from its solar orbit and put into one round the Earth. The drill-hole would then be plugged, restoring the stabilizing pressure. When the asteroid was called on to make the supreme sacrifice for its new primary, the hole would be unplugged again to steer it to its final, fatal encounter.

This elegant scheme can be put into effect long in advance of any threat. With the guard asteroid safely in orbit, the Earth is well defended. Of course, if the deadly incomer turns out to be another ice-cored asteroid, it might still be feasible to send a spacecraft to intercept it on the old plan. But instead of blowing it up, it could be steered into Earth orbit as an additional defender.

David Jones

The further Inventions of Daedalus (Oxford University Press), 148 past Daedalus columns expanded and illustrated, is now on sale. Special *Nature* offer: m.curtis@nature.com

Motherhood improves learning and memory

Neural activity in rats is enhanced by pregnancy and the demands of rearing offspring.

When a female mammal makes the transition from virginity to motherhood, she is forced to refocus her activities dramatically. She must adapt to a multitude of new demands by her offspring or risk losing a significant metabolic and genetic investment. She needs to find and remember the location of food stores, water sources and nest sites, and be able to exploit them to her offspring's advantage. The performance of these tasks may depend on a sharpening of her cognitive abilities. We have tested this idea by using two different strains of rat and two separate spatial tasks, and find that pregnancy and a litter of pups may combine to stimulate spatial learning and memory in females.

In the oestrous cycle, which occurs preparatory to mating, steroid hormones such as oestradiol and progesterone markedly increase the concentration of hippocampal apical dendritic spines in adult females, particularly when oestradiol levels are raised¹. Because dendritic spines increase the total surface area of the synapse, such neural changes may improve learning and memory². Pregnancy, which raises the amount of oestradiol and progesterone for significantly longer than the oestrous cycle, may produce long-lived improvements in behaviours that are not maternal as such, but which contribute to the survival and rearing of pups. The hormones of pregnancy may prepare sites that regulate learning and memory³, in much the same way that they stimulate the medial preoptic area to respond with maternal behaviour when pups are born⁴.

We investigated this idea by testing multiparous (animals that had given birth and lactated twice) and age-matched nulliparous (virgin) Sprague–Dawley females on a 16-day regimen of radial-arm maze tests. On the first six days, multiparous females made significantly more correct choices than nulliparous females ($P < 0.0002$ – 0.01 ; Fig. 1a).

In a second experiment, virgin female Long Evans rats were assigned to one of three groups: fosters, maternals and nulliparous. In the dry-land version of the Morris watermaze⁵, maternal females took significantly less time (43.2 s) than nulliparous rats (128 s) to recall and locate the food reward ($P < 0.05$; Fig. 1b). No significant differences were observed between foster (55 s) and maternal rats.

Taken together, these results show that a combination of reproductive and pup experience and stimulation is beneficial to learning and memory in female rats. Hormone-induced modifications to the hippocampus (increases in both long-term

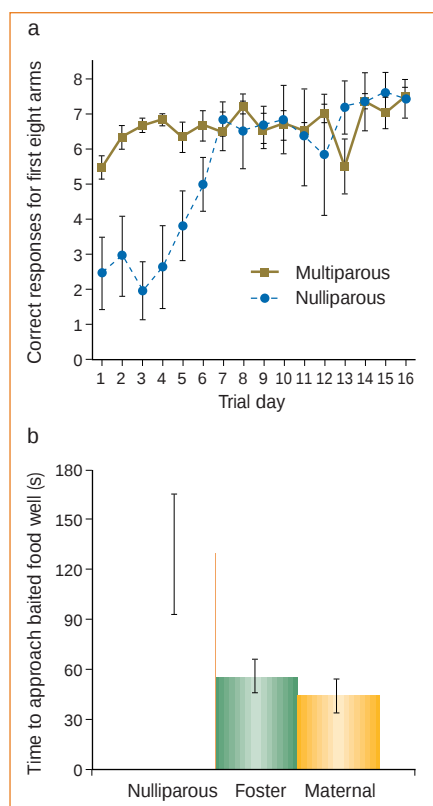


Figure 1 Improvement in spatial learning following reproductive and pup experience. **a**, Effects of distal reproductive experience on spatial learning. Two separate groups of 6–8 adult (~100 days old) randomly chosen female Sprague–Dawley rats were mated and allowed to deliver naturally. Litters were culled to 6 pups and allowed to remain with their mothers until weaning (21 days). Two weeks after weaning, these newly primiparous females were re-mated and the sequence repeated. These newly multiparous females were allowed to remain with their pups until weaning. Two weeks later, the multiparous females and their age-matched nulliparous cohorts were tested on the radial-arm maze. **b**, Effects of proximal reproductive experience on spatial learning in three groups of 5 virgin Long Evans rats: nulliparous rats, which had no mating or pup contact; fosters, for which females were exposed to three foster pups (exchanged every 24 h for milk-replete pups) for 16 days, when the females met all the criteria for displaying maternal behaviour (retrieving, grouping and crouching over pups within 60 min of exposure); and maternals, which are primiparous females that stayed with pups for 16 days. When pups were removed, females were weighed, food was restricted and behavioural testing began in a dry land maze⁵. Further experimental details are available from the authors.

potentiation⁶ and dendritic spine density^{1,7}) may improve the navigation skills involved in parental resource-gathering behaviours. For example, hippocampal volume in parental rodents during the breeding season is correlated with the size of the home range⁸. Concurrent alterations in both the hippocampus and the medial preoptic area may therefore contribute to the pronounced behavioural transition that is characteristic of the maternal female, with the medial preoptic area directly regulating maternal behaviour, and the hippocampus controlling supporting behaviours such as foraging.

As well as receiving the hormones of pregnancy, the female's brain is exposed to rich sensory events when pups are born. With each litter of pups comes a plethora of new sights, sounds, tastes, and tactile and suckling stimulation, the last of which can reorganize hypothalamic connections⁹. This sensory stimulation may have effects on brain structure and function that are similar to those caused by other types of enriched physical environment, an idea supported by the observation that the brains of parous, environmentally impoverished rats resemble those of rats in an enriched environment¹⁰.

Dendritic processes called filopodia are triggered by the excitatory activation of

N-methyl-D-aspartate (NMDA)¹¹, leading to the formation of synapses. Furthermore, new postsynaptic CA1 dendritic spines are rapidly formed after long-term potentiation (LTP, a learning analogue)¹². Neural activity brought about by pregnancy and the presence of pups may literally reshape the brain, fashioning a more complex organ that can accommodate an increasingly demanding environment.

Little attention has been paid to the remarkable neural plasticity that is inherent in reproduction itself and that underlies the subsequent behavioural changes, particularly those unique to late pregnancy and the postpartum period. To consider the relationship of a mother caring for her young as unidirectional disregards the potentially rich set of sensory cues in the opposite direction that can enrich the mother's environment. By providing such stimuli, the pups may ensure both their own and their mother's development and survival.

Craig H. Kinsley*, Lisa Madonia*, Gordon W. Gifford*, Kara Tureski*, Garrett R. Griffin*, Catherine Lowry†, Jamison Williams†, Jennifer Collins†, Heather McLearn†, Kelly G. Lambert†

*Department of Psychology, University of Richmond, Richmond, Virginia 23173, USA
e-mail: ckinsley@richmond.edu

†Department of Psychology,
Randolph-Macon College, Ashland,
Virginia 23005, USA

1. Woolley, C. S., Gould, E., Frankfurt, M. & McEwen, B. S. *J. Neurosci.* **10**, 4035–4039 (1990).
2. Daniel, J. M., Roberts, S. L. & Dohanich, G. P. *Physiol. Behav.* **66**, 11–20 (1999).
3. Loy, R., Gerlach, J. L. & McEwen, B. S. *Brain Res.* **467**, 245–251 (1988).
4. Numan, M. in *The Physiology of Reproduction* 2nd edn (eds Knobil, E. et al.) 221–302 (Raven, New York, 1994).
5. Kesner, R. & Dakis, M. *Psychopharmacology* **120**, 203–208 (1995).

6. Warren, S. G. et al. *Brain Res.* **703**, 26–30 (1995).
7. Kinsley, C. H. et al. *Soc. Neurosci. Abstr.* **24**, 952 (1998).
8. Jacobs, L. F., Gaulin, S. J. C., Sherry, D. F. & Hoffman, G. E. *Proc. Natl Acad. Sci. USA* **87**, 6349–6352 (1990).
9. Modney, B. K. & Hutton, G. I. *Mammalian Parenting: Biochemical, Neurobiological, and Behavioral Determinants* (eds Krasnegor, N. A. & Bridges, R. S.) 305–323 (Oxford Univ. Press, New York, 1990).
10. Diamond, M. C., Johnson, R. E. & Ingham, C. *Int. J. Neurosci.* **2**, 171–178 (1971).
11. Maletic-Savatic, R., Malinow, R. & Svoboda, K. *Science* **283**, 1923–1927 (1999).
12. Engert, F. & Bonhoeffer, T. *Nature* **399**, 66–70 (1999).

Laser optics

Fractal modes in unstable resonators

One of the simplest optical systems, consisting of two mirrors facing each other to form a resonator, turns out to have a surprising property. Stable resonators, in which the paths of the rays are confined between the two mirrors, have a well known mode structure (hermite–gaussian), but the nature of the modes that can occur in unstable resonant cavities (from which the rays ultimately escape) are harder to calculate, particularly for real three-dimensional situations¹. Here we show that these peculiar eigenmodes of unstable resonators are fractals, a finding that may lead to a better understanding of phenomena such as chaotic scattering and pattern formation. Our discovery may have practical application to lasers based on unstable resonators¹.

For a confocal unstable resonator consisting of a large concave mirror and a small convex feedback mirror sharing a common focus, the rays spill over the small mirror so that its size and shape define the output aperture (Fig. 1, insets); the other mirror is so large that its size and shape are irrelevant. The properties of an unstable resonator are determined by the round-trip magnification M and the (equivalent) Fresnel number N of the resonator: $N \equiv (M-1)a^2/2\lambda L$, where λ is the optical wavelength, $2a$ is the size of the small mirror, and L is the cavity length¹. The existence of a round-trip magnification M is characteristic for unstable resonators, as stable resonators lack such magnification. The value of M determines the round-trip spill-over loss around the small mirror, and N controls the size of the smallest details in the transverse mode profile as determined by diffraction. In an unstable-cavity laser, the losses are compensated by the gain provided by the laser medium.

Free-space propagation of the optical field is described by the Huygens–Fresnel integral¹, and cavity eigenmodes $u(x,y)$ are defined by the requirement that, after one round trip, the field profile remains unchanged apart from a complex multiplier

(the eigenvalue of the mode). We have calculated the mode patterns for a range of different aperture shapes, concentrating particularly on regular polygons and rhomboids. This calculation is numerically very demanding, so we used a non-orthogonal grid² to increase efficiency. For square or circular apertures, the diffraction problem can be reduced to a one-dimensional calculation, with the mode patterns for the square factorizing into x and y components.

A typical modal intensity distribution for triangular aperturing is shown in Fig. 1. Smaller and smaller triangles keep appearing as the mode pattern is repeatedly magnified. This shows the self-similar nature

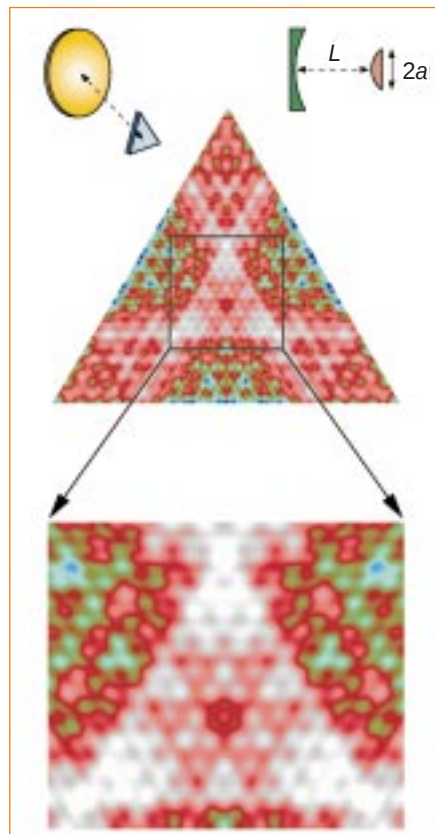


Figure 1 Calculated lowest-loss eigenmode of an unstable resonator with a triangular mirror. Left inset, a three-dimensional view; right inset, a side view. $M = 1.5$, $N = 10.5$. The intensity distribution $|u|^2$ on the small mirror is shown: white, high; red, medium; blue, low. Bottom, magnification of the centre part, with colour coding adjusted to exploit the full dynamical range.

of the mode, which continues down to the diffraction limit of approximately $a/4N$. On even smaller scales, further magnification does not reveal further substructure (Fig. 1, bottom). In the asymptotic limit $\lambda \rightarrow 0$ ($N \rightarrow \infty$), the diffractive spreading becomes negligible, leading to an ideal fractal. Large- N resonators are feasible (for example, $N = 14$ has already been realized in ref. 3).

The origin of the self-similarity can be understood intuitively: the existence of the round-trip magnification M means that the eigenmode must consist of (de)magnified copies of itself. Note that the hermite–gaussian modes of stable resonators are not self-similar because the round-trip magnification is absent.

A characteristic property of fractals is their non-integer fractal dimension D_f (refs 4,5). We found that an accurate determination of D_f for two-dimensional modes (as in Fig. 1) would require calculations at much higher N than considered here, and hence computational resources beyond those available; in earlier calculations⁶ for one-dimensional resonators with $N = 10^3$, we found that D_f was 1.6 ± 0.1 .

It has been shown that a tetraedical stacking of four reflecting spheres represents a chaotically scattering system with inherent fractal properties⁷. Our results help to explain this: light rays repeatedly scattering in the inner chamber of the tetraedical stacking undergo geometrical magnification owing to the reflecting spheres, just as in our unstable resonator. Coincidentally, the ‘aperture’ in ref. 7 is approximately triangular, like our aperture in Fig. 1, leading to similar patterns. The self-similar aspect seems to be a useful unifying feature of optical-pattern generation, providing insight into chaotic scattering⁷ and transverse nonlinear optics⁸. In the latter case, there is competition between bulk nonlinear behaviour and boundary-imposed linear diffraction as pattern-generating mechanisms.

G. P. Karman*†, G. S. McDonald*†, G. H. C. New†, J. P. Woerdman*

*Leiden University, PO Box 9504,
2300 RA Leiden, The Netherlands
e-mail: qo@molphys.leidenuniv.nl

†Blackett Laboratory, Imperial College,
London SW7 2BZ, UK

‡Present address: Philips Research Laboratories,
Prof. Holstlaan 4, 5656 AA Eindhoven,
The Netherlands

1. Siegman, A. E. *Lasers* Ch. 19, 22 (University Science, Mill Valley, California, 1986).
2. McDonald, G. S. & Firth, W. J. *J. Mod. Opt.* **40**, 23–32 (1993).
3. Cheng, Y.-J., Fanning, C. G. & Siegman, A. E. *Phys. Rev. Lett.* **77**, 627–630 (1996).
4. Falconer, K. *Fractal Geometry: Mathematical Foundations and Applications* (Wiley, New York, 1990).
5. Peitgen, H.-O., Jürgens, H. & Saupe, D. *Chaos and Fractals* (Springer, New York, 1992).
6. Karman, G. P. & Woerdman, J. P. *Opt. Lett.* **23**, 1909–1911 (1998).
7. Sweet, D., Ott, E. & Yorke, J. A. *Nature* **399**, 315–316 (1999).
8. Lugiato, L. A., Brambilla, M. & Gatti, A. *Adv. Atom. Mol. Opt. Phys.* **40**, 229–306 (1999).

Biodiversity

Seed germination in rainforest fragments

Habitat loss and fragmentation remain the greatest threats to the world's biodiversity¹. The local extinction of plant species from habitat fragments is common², although the reasons for this are not fully understood. Fragmentation is known to influence both birth- and death-related processes³, but the disruption of plant reproduction, especially pollination and seed production, is thought to be particularly important^{4,5}. The effects of fragmentation on post-pollination processes such as seed dispersal and germination have rarely been explored experimentally⁶. Here I show that seeds planted in forest fragments are less likely to germinate than those in continuous forest. This finding can have negative demographic consequences because it reduces the emergence of seedlings.

Heliconia acuminata is an Amazonian understory herb that is found in both the experimentally isolated rainforest fragments and the continuous forest reserves of Brazil's Biological Dynamics of Forest Fragments Project. During the 1998 fruiting season, which lasted from March to May, I collected 1,668 seeds from more than 200 plants in continuous forest. I then distributed these seeds to seven forest fragments (four fragments of one hectare and three of ten hectares) and three continuous forest sites. Seeds from each maternal plant were distributed evenly across sites to avoid confounding potential maternal effects with fragmentation-related effects.

In each site, seeds were planted every 10 metres along five 100-metre transects. At each location on the transect, a seed was placed on the ground, placed in a plastic cup planted into the ground and filled with local soil (preventing removal of seeds by ants), and placed in a cup buried in the same way but covered with a fine mesh (to prevent all seed predation and accumulation of leaf litter). Seeds were checked each month to see whether they had germinated. One year after planting, I compared the proportion of emerging seedlings in each treatment in each site by using two-way analysis of variance (ANOVA)⁷.

Seeds planted in continuous forest were between three and seven times more likely to germinate than those in forest fragments ($P < 0.0001$; Fig. 1). Seeds in fragments suffered from a variety of edge effects, including hotter, drier conditions and increased light penetration^{8,9}, all of which can affect the cues required for germination. I also found that seeds protected with mesh from leaf-litter accumulation and seed predation were more likely to germinate (although not significantly so), with the improvement being more marked in fragments than in

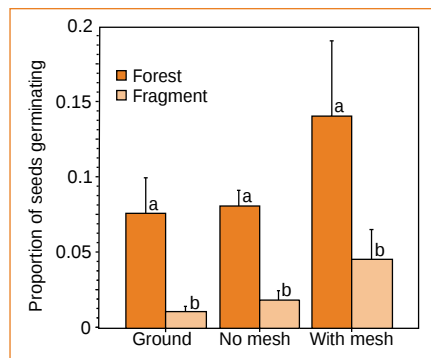


Figure 1 Mean proportion ($\pm$ s.e.m.) of seeds germinating in continuous forest and fragmented sites in different experimental treatments. Differences in the proportion germinating in each habitat and treatment were assessed by ANOVA. The square root of the percentage germinating in each plot was arcsine-transformed, and habitat type and treatment were treated as fixed effects. The effect of habitat type was highly significant ($F_{1,24} = 22.623$, $P < 0.0001$), but that of planting treatment was not ($F_{2,24} = 2.336$, $P = 0.118$). The planting treatment $\times$ habitat type interaction term was also nonsignificant ($F_{2,24} = 0.028$, $P = 0.973$). Bars with different letters are significantly different from each other.

forest sites (2.4- to 4.2-fold compared with 1.8- to 1.9-fold; Fig. 1). This asymmetrical improvement indicates that the greater accumulation of leaf litter in fragments (resulting from slower decomposition⁸) could be partly responsible for preventing germination as seeds become buried. Higher seed predation in fragments is an unlikely explanation, as predation is extremely low in all sites (E.M.B., unpublished data) and germination rates were similar for seeds placed on the ground and protected from predation in plastic cups.

The results could be explained by the adaptation of seeds to conditions in the area from which they were collected, with seeds failing to germinate in fragments because of a change in location, rather than as a direct result of fragmentation. However, this mechanism seems unlikely unless seeds are adapted to very broad conditions, such as 'continuous forest'. Seeds were collected from an area of more than 3.5 km², and one site of continuous forest was about 20 km from the collection area. When compared with other continuous forest sites, this site never had the lowest germination frequency for any treatment. Even if local adaptation were responsible for the observed pattern, the demographic consequences for *Heliconia* populations in these fragments would be the same. Populations in fragments have few flowering individuals (E.M.B. and W. J. Kress, unpublished data), so most seeds for these populations probably come from outside the fragments.

Seeds from tropical rainforest plants rarely survive in seed banks for more than a year because they are highly susceptible to fungal pathogens, seed predation and burial under leaf litter¹⁰. My results therefore

represent the outcome of a population's entire reproductive effort for a year. Although germination success is likely to vary from year to year, my results indicate that seeds dispersed into rainforest fragments are less likely to germinate than those in continuous forest.

Furthermore, my estimates of the effect of fragmentation on seed germination are probably conservative, as plants reproducing in forest fragments may be inbred and suffer from reduced heterozygosity¹¹. If inbred seeds are less likely to germinate, these genetic effects could further reduce the recruitment of seedlings into forest fragment populations, helping to explain why plant populations in habitat fragments often fail to persist in the long term.

Emilio M. Bruna

Department of Environmental Science and Policy, and Center for Population Biology, University of California, Davis, California 95616, USA and Biological Dynamics of Forest Fragments Project, National Institute for Research in the Amazon (INPA), CP 478, Manaus AM 69011-970, Brazil
e-mail: embruna@ucdavis.edu

1. Wilcove, D. S., Rothstein, D., Dubow, J., Phillips, A. & Losos, E. *Bioscience* **48**, 607–615 (1998).
2. Turner, I. M., Chua, K. S., Ong, J. S. Y., Soong, B. C. & Tan, H. T. W. *Conserv. Biol.* **10**, 1229–1244 (1995).
3. Jules, E. S. *Ecology* **79**, 1645–1656 (1998).
4. Bond, W. J. in *Extinction Rates* (eds Lawton, J. H. & May, R. M.) 131–146 (Oxford Univ. Press, New York, 1995).
5. Aizen, M. A. & Feinsinger, P. *Ecology* **75**, 330–351 (1994).
6. Wong, T. C. M., Sodhi, N. S. & Turner, I. M. *Biol. Conserv.* **85**, 97–104 (1998).
7. Sokal, R. R. & Rohlf, F. J. *Biometry* (Freeman, New York, 1995).
8. Didham, R. K. *Oecologia* **116**, 397–406 (1998).
9. Kapos, V., Wandelli, E., Camargo, J. L. & Ganade, G. in *Tropical Forest Remnants: Ecology, Management, and Conservation of Fragmented Communities* (eds Laurance, W. F. & Bierregaard, R. O. Jr) 33–44 (Univ. Chicago Press, 1997).
10. Vázquez-Yanes, C. & Orozco-Segovia, A. *Annu. Rev. Ecol. Syst.* **24**, 69–87 (1993).
11. Young, A., Boyle, T. & Brown, T. *Trends Ecol. Evol.* **11**, 413–418 (1996).

Vision

Dichromatism in macaque monkeys

Old World primates have trichromatic vision because they have three types of cone photoreceptor, each of which is maximally sensitive to short, middle or long wavelengths of light¹. Although a proportion of human males (about 8% of caucasians, for example) have X-chromosome-linked colour-vision abnormalities², no non-human Old World primates have been found to be colour-vision defective^{3,4}. We have tested 3,153 macaque monkeys but found only three dichromats, a frequency that is much lower than in humans.

Most defects in colour vision in humans arise from the loss of the gene that encodes

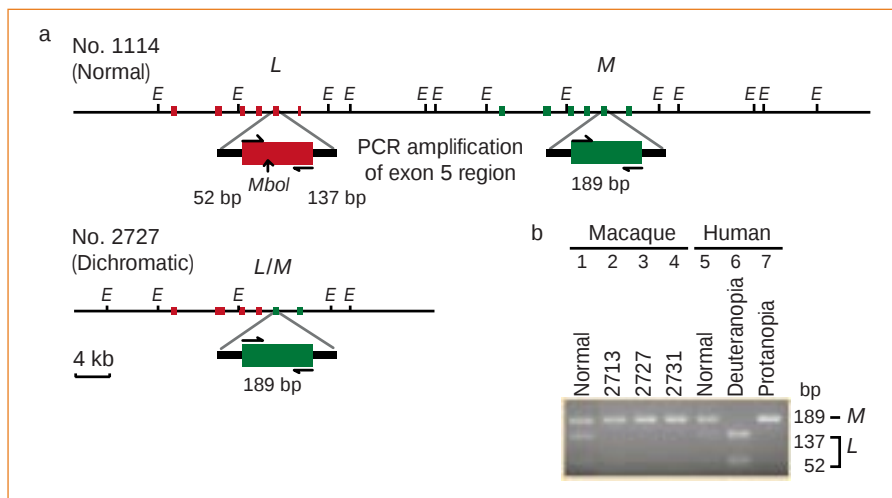


Figure 1 Physical maps of the *L* and *M* genes. **a**, Top, the 77-kilobase (kb) region of a wild-type macaque (no. 1114), showing exons of the *L* (red) and *M* (green) genes. Enlarged exon-5 regions are shown underneath. The PCR primers and products for exon 5 of the *L* and *M* genes are shown in red (137 and 52 base pairs (bp), cleaved by *MboI* digestion) and green boxes (189 bp, uncleaved), respectively. An *MboI* site specific for exon 5 of the *L* gene is conserved among Old World monkeys⁸. Bottom, the 36-kb region of a dichromatic candidate (no. 2727). The enlarged PCR-amplified fragment is shown underneath. **b**, Gel electrophoresis analysis. The bands of 189, 137 and 52 bp shown in **a** are indicated as *M* (green) or *L* (red). Lane 1, macaque wild type; lanes 2–4, subjects 2713, 2727 and 2731, respectively, showing only the exon 5 fragment of the *M* gene; lane 5, human wild type; lanes 6 and 7, deuteranopic (*M* cone absent) and protanopic (*L* cone absent) humans, respectively.

pigments sensitive to either middle (*M*) or long (*L*) wavelengths of light. These genes are located on the X chromosome at Xq28 in a head-to-tail tandem array, and the repetitive units are gained or lost, and hybrid pigment genes generated, by unequal chromosomal recombination^{5–7}.

In our attempt to find dichromatic monkeys, we co-amplified the exon-5 region of the *L* and *M* genes with a set of common primers and discriminated between them with an *MboI* restriction site specific for the exon-5 fragment of the *L* gene⁸ (Fig. 1a). We studied DNA from 744 male and 1,301 female crab-eating monkeys and 455 males and 653 females from 18 other macaque species, including samples from the Tsukuba Primate Center and the Nihon Monkey Center, and found only three DNA samples from males that lacked the exon-5 fragments of the *L* gene (Fig. 1b). These were from groups of crab-eating monkeys (26 males and 31 females) whose blood was sampled in the Pangandaran National Park, Indonesia.

The frequency of the dichromat males in this species (3 out of 744) is 0.4%, a value that is significantly lower (chi-squared test, $P < 0.002$) than in human males, for which a survey of 30,000 caucasian males found that 2% were dichromats². No dichromat candidates were found in other macaque species.

We then constructed physical maps of the *L* and *M* genes from a wild-type (no. 1114) and a dichromatic (no. 2727) monkey from two genomic DNA libraries. The *L* and *M* genes were both present in the wild type but the dichromatic monkey had only a single pigment gene (Fig. 1a).

Nucleotide sequence analysis then revealed that this single pigment gene was an *L/M* hybrid (called *R4G5*) consisting of exons 1 to 4 from *L* and exons 5 and 6 from *M*, indicating that the recombination point was located between exon 4 of *L* and exon 5 of *M*. Southern hybridization analysis revealed the presence of the *L/M* hybrid gene but no other *L* and *M* genes in this monkey. We found by polymerase chain reaction (PCR) using primers that specifically detect the *R4G5* hybrid gene that these three males and two females carried the hybrid gene.

We then used transient transfection into 293T cells to express complementary DNAs encoding the *M*, *L* and hybrid (*R4G5*) pigments (Fig. 2). Analysis of photobleaching difference absorption spectra showed that the absorbance maxima ($\lambda_{\max}$) of reconstituted *L*, *M* and hybrid pigments are 564, 532 and 538 nm, respectively. The $\lambda_{\max}$ for the hybrid was red-shifted by 6 nm from that of the *M* pigment, indicating that these monkeys are almost protanopic (insensitive to red)^{9,10}.

In humans, between 58% and 78% of individuals possess several *M* genes^{11,12}, perhaps as a result of unequal recombination. In contrast, examination of 129 male crab-eating monkeys by quantitative Southern hybridization revealed that only six monkeys (5%) have more than one *M* gene, up to a maximum of four. All the other monkeys, apart from the dichromats, have just one *L* and one *M* gene. The variation in the number of *L* and *M* genes may be rare in macaques, but the sequence of the visual pigment genes is similar to that in humans^{5,6}.

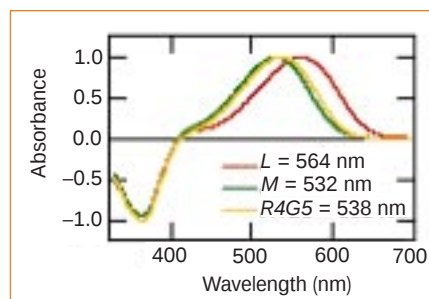


Figure 2 Absorbance spectra of *L*, *M* and hybrid (*R4G5*) pigments. Pigments expressed in 293T cells were extracted from the membrane fractions after 11-*cis*-retinal was added. Difference spectra were calculated from the spectra before and after photobleaching with light of wavelength ≥ 590 nm for 1 min in the presence of 10 mM NH_4OH (ref. 13). The $\lambda_{\max}$ values of the *L* and *M* pigments are in good agreement with those previously measured by microscopic photometry¹⁴.

Akishi Onishi*†, Satoshi Koike*‡, Miki Ida*‡, Hiroo Imai†, Yoshinori Shichida†, Osamu Takenaka§, Akitoshi Hanazawa||, Hidehiko Konatsu||, Akichika Mikami¶, Shunji Goto#, Bambang Suryobroto☆, Kenji Kitahara, Tetsuo Yamamori***

*Laboratory for Speciation Mechanisms I, National Institute for Basic Biology, 38 Myodaijicho, Okazaki 444-8585, Japan
e-mail: yamamori@nibb.ac.jp

†Department of Biophysics, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan
‡Department of Microbiology and Immunology, Tokyo Metropolitan Institute for Neuroscience, Fuchu, Tokyo 183-8526, Japan

§Departments of Cellular and Molecular Biology and ¶Behavioural and Brain Sciences, and #Center for Applied Primatology and Human Evolutionary Modelling Research, Primate Research Institute, Kanrin, Inuyama, Aichi 484-8506, Japan

||Laboratory of Neural Control, National Institute for Physiological Sciences, Myodaijicho, Okazaki 444-8585, Japan

☆Department of Biology, Bogor Agricultural University, Jalan Pajajaran, 16143 Bogor, Indonesia

**Department of Ophthalmology, Jikei University School of Medicine, Minatoku, Tokyo 105-0003, Japan

- Tovée, M. J. *Trends Neurosci.* **17**, 30–37 (1994).
- Fletcher, R. & Voke, J. *Defective Colour Vision: Fundamentals, Diagnosis and Management* (Hilger, Bristol, 1985).
- Jacobs, G. H. & Harwerth, R. S. *Am. J. Primatol.* **18**, 35–44 (1989).
- Jacobs, G. H. & Deegan, J. F. *Vis. Neurosci.* **14**, 921–928 (1997).
- Nathans, J., Thomas, D. & Hogness, D. S. *Science* **232**, 193–202 (1986).
- Vollrath, D., Nathans, J. & Davis, R. W. *Science* **240**, 1669–1672 (1988).
- Deeb, S. S. et al. *Am. J. Hum. Genet.* **51**, 687–700 (1992).
- Dulai, K. S. et al. *Vision Res.* **34**, 2483–2491 (1994).
- Merbs, S. L. & Nathans, J. *Science* **258**, 464–466 (1992).
- Neitz, J., Neitz, M. & Kainz, P. N. *Science* **274**, 801–804 (1996).
- Drummond-Borg, N., Deeb, S. S. & Motulsky, A. G. *Proc. Natl Acad. Sci. USA* **86**, 983–987 (1989).
- Jorgense, A. L., Deeb, S. S. & Motulsky, A. G. *Proc. Natl Acad. Sci. USA* **87**, 6512–6516 (1990).
- Kojima, D. et al. *Biochemistry* **35**, 2625–2629 (1996).
- Bowmaker, J. K., Dartnall, H. J. & Mollon, J. D. *J. Physiol. (Lond.)* **298**, 131–148 (1980).

A protein kinase encoded by the *t* complex responder gene causes non-mendelian inheritance

Bernhard G. Herrmann, Birgit Koschorz, Karin Wertz, K. John McLaughlin* & Andreas Kispert

Max-Planck-Institute of Immunobiology, Stübeweg 51, 79108 Freiburg, Germany

Males heterozygous for the *t*-haplotype form of mouse chromosome 17 preferentially transmit the *t*-chromosome to their progeny. Several *distorter/sterility* loci carried on the *t*-haplotype together impair flagellar function in all spermatozoa whereas the *responder*, *Tcr*, rescues *t*-sperm but not wild-type sperm. Thus, *t*-sperm have an advantage over wild-type sperm in fertilizing egg cells. We have isolated *Tcr* by positional cloning and show that it is a member of a novel protein kinase gene family, designated *Smok*, which is expressed late during spermiogenesis. *Smok* kinases are components of a signal cascade which may control sperm motility. *Tcr* has a reduced kinase activity, which may allow it to counterbalance a signalling impairment caused by the *distorter/sterility* loci. *Tcr* transgene constructs cause non-mendelian transmission of chromosomes on which they are carried, which leads to sex-ratio distortion when *Tcr* cosegregates with the Y chromosome.

Transmission ratio distortion (TRD) in mice is a phenomenon associated with the *t*-complex, a 12-cM region in the proximal part of chromosome 17 (refs 1, 2). In wild mouse populations the *t*-complex exists in two forms: the wild-type and the *t*-haplotype. Males heterozygous for a complete *t*-haplotype (*t*/+) preferentially transmit the *t*-haplotype chromosome. Meiotic recombination between wild-type and *t*-haplotype DNA is suppressed, allowing the *t*-haplotype to be transmitted *en bloc* to the offspring. However, so-called partial *t*-haplotypes carrying only parts of the *t*-complex in the *t*-form have been isolated. They contain only a subset of the factors involved in TRD and have allowed the identification and mapping of six distinct loci acting together to achieve a high transmission ratio^{3,4}.

Factors involved in TRD are expressed during spermatogenesis and affect spermatozoa. The *distorters* (*Tcd1* to *Tcd5*) act in all sperm cells, whereas the *responder* (*Tcr*) acts only in sperm carrying the *Tcr* locus³. Homozygosity at regions carrying *distorter* loci results in male sterility, indicating that TRD and sterility may be caused by the same mutant loci⁵.

TRD and male sterility have been formally explained in a model⁶ as follows: heterozygosity at *distorter* loci produces a strong and harmful effect on the wild-type form of the *responder*, whereas its effect on the *t-responder* (*Tcr*) is weak, leading to distortion in favour of *Tcr*. Homozygosity at *distorter* loci, however, has a strong harmful effect on both the wild-type *responder* and *Tcr*, leading to sterility.

It has been shown that *t*-haplotypes compromise sperm flagellar function⁶. Sperm from males heterozygous for complete *t*-haplotypes exhibit vigorous but non-progressive motility. This behaviour affects their straightness of path and may hinder sperm in reaching the site of fertilization. These abnormalities are presumably associated with the wild-type sperm, leading to an advantage for *t*-sperm in reaching the eggs. Abnormalities of flagellar movement are enhanced in sperm from males homozygous for complete *t*-haplotypes, which are sterile, suggesting that *in vivo* such sperm cannot reach the site of fertilization⁶. Thus, TRD and male sterility may be caused by mutations in factors involved in the control of flagellar movement. *Tcr* has been mapped to *Leh66b*, a subregion of a locally

dispersed family of large DNA elements in the centre of the *t*-complex⁷⁻⁹. Previous studies have localized *Tcr* to a region with a maximum size of 155 kilobases (kb), defined by the recombination breakpoints of the partial *t*-haplotypes *t*^{w71/r1}, which does not have *Tcr* activity, and *t*^{h49}, which contains *Tcr* (ref. 10). Attempts to isolate *Tcr* have led to the cloning and molecular analysis of a candidate gene, *Tcp10b*¹ (ref. 11). However, mutagenesis of *Tcp10b*¹ by homologous recombination showed that it is not the *responder*¹².

Tcr defines a new protein kinase family

A large part of the region known to contain *Tcr* was isolated from the *t*-haplotype *t*^{w12} (Fig. 1a). It contains a fragment of a ribosome *S6* kinase 3 (*Rsk3*) gene which lacks several exons at the 5' end¹³. This gene is expressed during spermatogenesis (Fig. 1c), which suggested that its 5' end is derived from another gene. Extensive searches for the 5' end led to the isolation of a complementary DNA clone, 66bk/4-2, representing a fusion transcript consisting of the 5' untranslated region and open reading frame (ORF) of a novel protein kinase gene and *Rsk3* sequences (Fig. 1a). The ORF encodes 484 amino-acid residues and translates into a protein with a relative molecular mass of 54,626 (*M_r* ≈ 55K). It shows highest similarity to the MARK Ser/Thr protein kinase family, which phosphorylates microtubule-associated proteins and triggers microtubule disruption¹⁴. We suspected that this fusion gene is *Tcr*; the subsequent genetic and functional analyses confirmed this hypothesis.

Southern blot analysis of genomic mouse DNA revealed a number of *Bam*HI fragments hybridizing to the novel kinase gene, indicating that *Tcr* may be a member of a gene family (Fig. 1b). Two hybridizing fragments (B7.8 and B9.1) were assigned to *Leh66b*. B7.8 contains the ORF of *Tcr* and B9.1 that of another family member. The latter gene and the family were designated 'sperm motility kinase' (*Smok*).

The close proximity and sequence relationship of *Tcr* and *Smok* indicate that they may have arisen by duplication of a common ancestor. *Smok* kept the original gene structure, whereas *Tcr* derived from a rearrangement with the neighbouring *Rsk3* gene. Additional amino-acid modifications produced a dominant factor. *Tcr* thus represents a mutant form of *Smok*. In accordance with nomenclature rules it is abbreviated by the symbol *Smok*^{Tcr}. However, to make it easier to distinguish between wild-type and mutant forms we use the original symbol *Tcr* here.

* Present address: Center for Animal Transgenesis and Germ Cell Research, The School of Veterinary Medicine, New Bolton Center, Myrin Building, 382 W. Street Rd, Kennett Square, Pennsylvania 19348, USA.

Smok genes have probably been amplified in association with *Leh66* elements⁸, raising the question of whether the fusion gene was also amplified. Analysis of testis RNA by polymerase chain reaction with reverse transcription (RT-PCR) showed that *Tcr* gene transcripts are exclusively derived from the gene identified in *Leh66b*, indicating that this gene may be unique (Fig. 1c).

Several members of the *Smok* gene family were isolated. An amino-acid sequence comparison of putative *Smok* proteins with MARK2 is shown in relation to the *t*⁶ allele of *Tcr* (Fig. 2). Within the amino-terminal protein kinase catalytic domain of *Tcr*, 95 residues out of 249 (38%) are identical to MARK2, whereas in the carboxy-terminal region the two proteins appear to be unrelated (not shown). All *Smok* sequences and *Tcr* are highly related to each other. However, *Tcr* diverges in a number of residues from all other *Smok* protein sequences. The ORFs of two family members extend by 20 additional codons at the N terminus, thus representing a subfamily, designated *Smok2*.

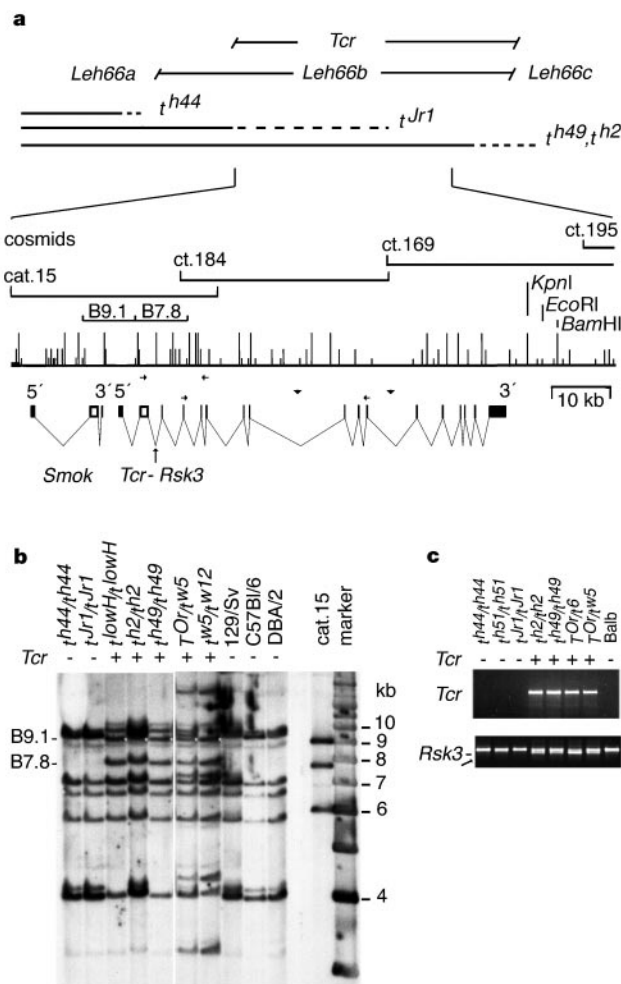


Figure 1 Positional cloning identifies *Tcr* as a member of a novel gene family. **a**, Genetic and physical map of the region containing *Tcr*. *Tcr* arose by fusion of a *Smok* gene which lost its 3' end with an *Rsk3* gene which lost its 5' end. A fusion transcript is formed by RNA splicing (vertical arrow). Exons containing the ORFs of *Smok* and *Tcr* are shown by open boxes; triangles indicate exons missing in *Rsk3*; horizontal arrows indicate PCR primer pairs (upper 181/144, lower 145/146). **b**, Southern blot hybridization using exon 1/exon 2 of *Tcr* as probe detects many members of the *Smok* gene family. Two fragments (B7.8 and B9.1, marked by asterisks) can be assigned to the region containing *Tcr*. **c**, RT-PCR experiment detecting transcripts derived from the *Tcr*-*Rsk3* fusion gene. Wild-type *Rsk3* transcripts are detectable in all testis RNA samples analysed (primer pair 145/146); an additional smaller fragment resulting from a deletion of one exon in the *Rsk3* gene in *Leh66b* (arrow) and the *Tcr*-*Rsk3* fusion cDNA fragment (primer pair 181/144) are detected only in RNA from males carrying *Tcr*.

The catalytic domains of both *Smok*⁺¹²⁹ and *Tcr* function as kinases *in vitro* (Fig. 3d). However, on the artificial substrate myelin basic protein the kinase activity of *Tcr* is an order of magnitude lower than that of *Smok*⁺¹²⁹, suggesting that the *Tcr* kinase activity may also be reduced significantly on native substrates. On the other hand, the autophosphorylation activity of *Tcr* is only two-fold lower than that of *Smok*, indicating that one of two autophosphorylation sites may be missing in *Tcr*. One of these sites may be threonine 164 (T164), which is replaced by an isoleucine in *Tcr*. T164 is located in subdomain VIII, which is involved in the activation or upregulation of many protein kinases¹⁵ including MARK2 (ref. 14). Another alteration of *Tcr* which may contribute to its low kinase activity is T129, which replaced a highly conserved lysine residue in subdomain VIIb, the kinase active site¹⁵. This lysine appears to be required for neutralizing the negative charge of the γ -phosphate during transfer¹⁵.

Expression of *Tcr* and *Smok* late in spermiogenesis

The RNA expression of *Smok* genes and *Tcr* was investigated by northern blot analysis, polymerase chain reaction with reverse transcription (RT-PCR) and *in situ* hybridization. Northern blot analysis showed that *Smok* genes are expressed in the testis from 22 days post partum. None of the other tissues tested was positive, indicating that *Smok* gene expression may be highly tissue specific. A transcript derived from *Tcr*, expected to be ~7.2 kb long, could not be detected by northern analysis, but RT-PCR experiments showed that *Tcr* transcription starts at the same stage as *Smok* gene transcription (Fig. 3b).

Spermatogenesis starts soon after birth, and approximately 35 days are necessary to produce fertile spermatozoa¹⁶. In testes dissected from prepubertal males, progressively more advanced stages of spermatogenesis are observed. In the testis of 22-day-old mice, spermiogenesis (the transformation of round haploid spermatids into motile spermatozoa) is in progress. It has been shown that *protamine-1* (*Prm1*) is transcribed in round spermatids¹⁷, whereas the testis-specific expression of *angiotensin converting enzyme* (*Ace*)¹⁸ and *Kit*¹⁹ is observed in elongating spermatids. Using RT-PCR analysis, we detected *Prm1* transcription from 18 days post partum, and *Ace* and *Kit* transcription from 22 days post partum onwards. The expression profiles of *Ace* and *Kit* coincide with those of *Smok* genes and *Tcr*, indicating that transcription of the latter may also start in elongating spermatids. *In situ* hybridization experiments on testicular sections confirmed these results by revealing transcripts of *Smok/Tcr* genes in cells in the central portion of the seminiferous tubule adjacent to the lumen, in elongating spermatids and in spermatozoa (Fig. 3c).

The late expression of *Smok* genes and *Tcr* during spermiogenesis supports genetic data suggesting haploid action of *Tcr*. However, as spermatids at that stage are still connected by intercellular bridges which may allow the passage of soluble components^{20,21}, it is not clear how *Smok/Tcr* may be retained in the cells producing them. Translational control of transcripts and immobilization of protein products in subcellular compartments may be involved. In any case, the haploid action of *Tcr* indicates that each spermatozoon may produce its own set of *Smok/Tcr* protein products. This might create functional differences between individual spermatozoa.

Tcr causes non-mendelian transmission of chromosomes

Tcr has been identified as a dominant factor promoting its own propagation with the help of *distorter* loci. If current views about *Tcr* function are correct, *Tcr* should also promote TRD for any other chromosome into which it is integrated. To test this assumption, two transgenic constructs containing the ORF of *Tcr* were randomly integrated into the genome. In one construct (*Tg4*), the testis-specific promoter of *Kit*¹⁹ was used to drive *Tcr* expression in elongating spermatids. The other construct (*Tg5*) carries exon 1 and upstream sequences of *Tcr*^{rw12}, representing the presumptive

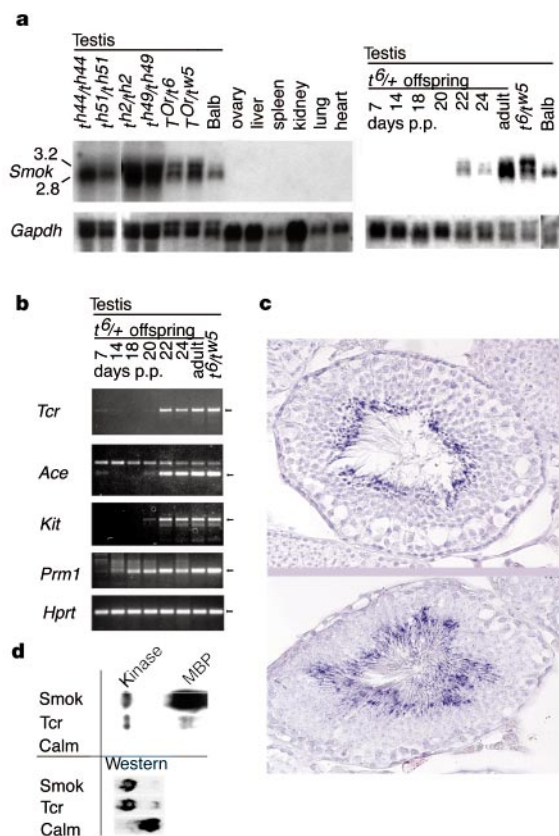


Figure 3 *Smok* genes and *Tcr* are expressed during spermiogenesis. **a**, Northern blot analysis using exon 1/exon 2 of *Tcr* as probe. **b**, RT-PCR analysis of *Tcr*, *Ace*, *Kit*, *Prm1* and *Hprt* transcription in the testis. Days p.p., days post-partum. **c**, *In situ* hybridization of testis sections showing *Smok/Tcr* gene transcripts in elongating spermatids and spermatozoa. **d**, Protein kinase assay of *Smok*⁺¹²⁹ (*Smok*) and *Tcr*¹⁶ (*Tcr*). Both phosphorylate myelin basic protein (MBP) and autophosphorylate (kinase), but their kinase activities are significantly different. Calmodulin (Calm) served as negative control. The relative amounts of protein used in the assay were controlled by western blot analysis.

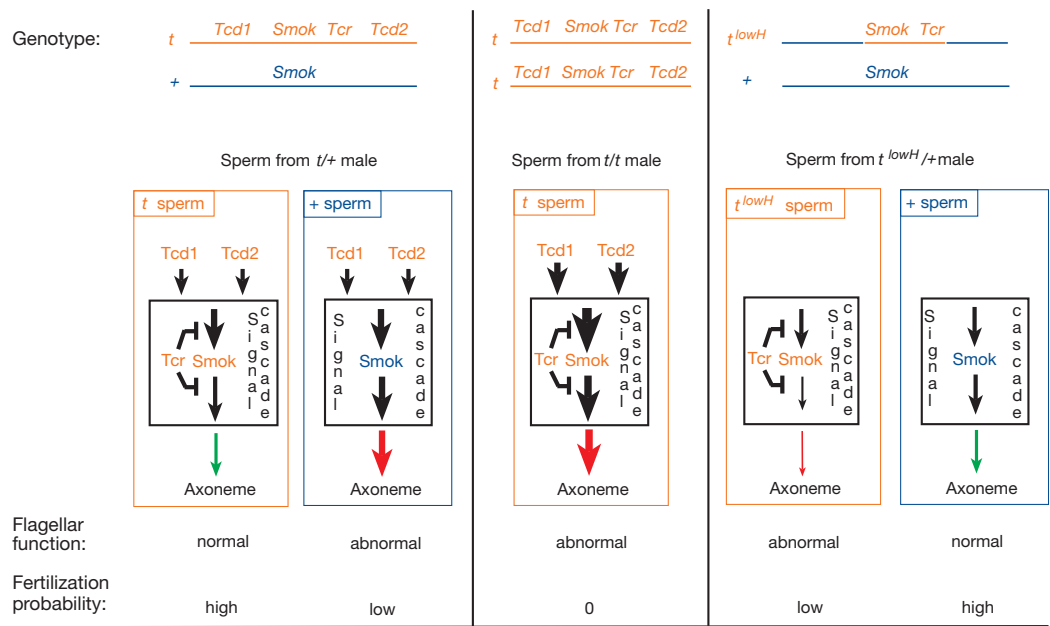


Figure 4 Proposed mechanism producing TRD and male sterility. The model is described in the text. For simplification, only the strong *t-distorter/sterility* loci *Tcd1/tcs1* (*Tcd1*) and *Tcd2/tcs2* (*Tcd2*) and their products are indicated, whereas their wild-type alleles and

restore normal downstream signalling, resulting in the rescue of normal flagellar movement.

As *Tcr* and *Smok* are co-expressed and act in parallel in the same sperm cells, they must be competing for interacting partners. It is expected that the C-terminal domain might be important in such interactions, and that it might be functionally different between *Tcr* and *Smok*. The effect of *Tcr* on sperm cells would thus depend on the *Tcr/Smok* dosage ratio, the affinity of either protein for binding sites/interaction partners, the catalytic activities and as-yet unknown properties of *Tcr/Smok* proteins. In other words, the better *Tcr* can outcompete *Smok*, the stronger the effect on spermatozoa.

Transmission at non-mendelian ratios occurs when the signalling activity of *Smok/Tcr* results in abnormal flagellar function in one type of sperm, while other sperm cells are normal. Enhanced *Smok* signalling may occur in wild-type sperm derived from *t/+* males expressing *t-distorter/sterility* loci, while *t*-sperm are rescued by *Tcr* (Fig. 4, left). Reduced *Smok/Tcr* signalling, on the other hand, may occur in *t*^{lowH}-sperm or *Tcr*-transgenic sperm expressing *Tcr* on a wild-type background (Fig. 4, right). In the former case, transmission ratio distortion occurs in favour of *t*-sperm; in the latter case, in favour of wild-type sperm. Homozygosity at the *t-distorter/sterility* loci cannot be rescued by *Tcr* and results in severe sperm abnormality and male sterility (Fig. 4, middle).

It has been shown that *t*-haplotypes behave differently on various genetic backgrounds²² and that TRD also occurs in males heterozygous for two different *t*-haplotypes²³. These observations may be explained by allelic variability in *Smok* and *Tcr* genes carried by different mouse inbred strains and different *t*-haplotypes. Thus, TRD may also occur in F1 males derived from inbred lines expressing *Smok* proteins with significantly different kinase activity. TRD or male sterility may also occur in other species, including humans, provided that *Smok* genes are functionally conserved, as a result of mutations in *Smok* genes affecting sperm motility. Low fertility of some mouse inbred lines and some forms of infertility observed in humans may be due to impairment of *Smok* function.

The investigation of factors controlling sperm flagellar movement has led to the isolation of a number of proteins involved in signal cascades activated by Ca²⁺ or cyclic AMP^{24–27}. Controlled phosphorylation and dephosphorylation of axonemal components are

products are not. *t*-haplotype alleles and factors are shown in orange, wild-type in blue. Arrows indicate signalling (green, normal; red, abnormal), bars indicate counteraction; widths of arrows or bars symbolize strength of effect.

Table 1 Transmission of Tg construct or sex from males carrying a Tcr transgene

			Offspring			
Transgene	Genotype of male	Number of males	Tg	Non-Tg	Total	% Tg
<i>Tg5-43</i>	<i>t^{h51}t^{h18}/tf</i>	10	316	103	419	75.4
<i>Tg5-43</i>	<i>tf/B6</i>	8	233	215	448	52
<i>Tg5-25</i>	<i>t^{h51}t^{h18}/tf</i>	4	186	32	218	85.3
<i>Tg5-25</i>	<i>tf/tf</i>	5	118	205	323	36.5
			Males	Females	Total	% male
<i>Tg4-3</i>	<i>t^{h51}t^{h18}/tf</i>	7	217	114	331	65.5
<i>Tg4-3</i>	<i>tf/tf</i>	5	231	240	471	49
		χ^2	<i>P</i>			
<i>Tg5-43/0</i> ; <i>tf/B6</i>	vs	<i>Tg5-43/0</i> ; <i>t^{h51}t^{h18}/tf</i>	50.07	<0.001		
<i>Tg5-25/0</i> ; <i>tf/tf</i>	vs	<i>Tg5-25/0</i> ; <i>t^{h51}t^{h18}/tf</i>	123.88	<0.001		
<i>Tg4-3/0</i> ; <i>tf/tf</i>	vs	<i>Tg4-3/0</i> ; <i>t^{h51}t^{h18}/tf</i>	20.83	<0.001		

Tg, transgenic.

essential for controlling flagellar movement^{28,29}. Therefore, it is likely that Smok/Tcr and the distorter/sterility factors are components of Ca²⁺- and/or cAMP-dependent signal cascades.

Two testis-expressed genes isolated from the *t*-complex, *Tctex1* and *Tctex2*, have been suggested as candidates for *tcs1* and *Tcd3/tcs3*, respectively^{30,31}. Homologues of *Tctex1* and *Tctex2* isolated from *Chlamydomonas* encode light-chain components of the inner and outer dynein arms of the axoneme, respectively^{32,33}. This indicates that either or both proteins may act downstream of Smok kinase. Tcpl1, a proposed cell-surface receptor involved in modulating the adenylyl cyclase/cAMP pathway, has been proposed as a candidate for the distorter locus *Tcd2* (refs 2, 34). An involvement of Tcpl1 in ratio distortion remains to be demonstrated. *Tcr* is the first mammalian gene isolated for which a function in segregation distortion has been shown experimentally. This opens the way for the molecular investigation of TRD. □

Methods

Mice

The partial *t*-haplotypes t^{h44} , t^{h51} and $t^{w71/r1}$ (t^{r1}) do not carry *Tcr*, whereas t^{lowH} , t^{h2} , t^{h49} , t^6 and the complete *t*-haplotypes t^{w5} and t^{w12} do^{3,9,10}. The *t*-haplotype $t^{h51}t^{h18}$ carries the *distorter/sterility* loci *Tcd1/tcs1* and *Tcd2/tcs2*, but not *Tcr*; *tf* is a recessive mutation marking the wild-type chromosome used in studies describing TRD³.

Transgene constructs were injected into eggs derived from matings of BCB females with NMRI males. Transgenic founders were crossed to tf/tf or $t^{h51}t^{h18}/tf$ mice and male offspring of the genotype $t^{h51}t^{h18}/+$; Tg/0 or $+/tf$; Tg/0 were then crossed to $t^{h51}t^{h18}/tf$ or tf/tf females. Males of the genotype $t^{h51}t^{h18}/tf$; Tg/0, $+/tf$; Tg/0 or tf/tf ; Tg/0; were mated to NMRI females. Offspring were tested for transgene carriers, either by PCR analysis of embryos or by scoring external genitalia of newborn pups. Genotyping of males and offspring for t^{h51} , t^{h18} , *tf* or other wild-type chromosomes was done by PCR testing for the markers *Tcp1*, *Hba-4ps* and *D17Mit46*, respectively.

Genomics

The isolation of cosmids from *Leh66b* and characterization of an *Rsk3* gene fragment located in this region have been described¹¹. We isolated the cDNA clone 66bk/4-2 encoding Tcr from a testis cDNA library. It is 2,807 base pairs (bp) long and includes part of exon 1 (bp 1–437), a long ORF (bp 524–1,975) encoded on exon 2 (bp 438–2,023) and *Rsk3*-derived sequences (2,024–2,807) of *Tcr*¹⁶, but does not represent all exons of *Rsk3*. Tcr alleles of t^{w5} , t^{w12} and t^6 , t^{h2} , t^{h49} differ by one amino acid (Fig. 2), in agreement with genetic data³⁵ and further supporting the identification of the cloned gene reported here as *responder*. Hybridization of the genomic Southern blot shown in Fig. 1b was done with a probe covering bp 190–1,884 of 66bk/4-2. The sequences of *Smok*^{w12} and *Smok*⁺¹²⁹ are derived from genomic clones, *Smok2*^{w5} from a cDNA clone, and *Smok2*^{+Balt} from a cDNA derived by RT-PCR of testis RNA. Testis cDNA libraries were prepared using the SuperScript plasmid cDNA cloning system (Life Technologies) and screened using standard or PCR-based techniques. RT-PCR was carried out using SuperScript II (Life Technologies). DNA was sequenced using fluorescently labelled dye-terminator-based sequencing methods on an ABI Prism 310 Genetic Analyzer (PE Applied Biosystems) and analysed using the MacMolly Tetra software package (Soft Gene).

Transcript analysis

RNA was isolated and analysed by northern blot hybridization (20 µg total RNA per lane) as described³⁶ using a probe covering bp 188–1,884 of 66bk/4-2, or by RT-PCR. The following primer pairs were used in RT-PCR analyses assaying testis-specific transcription

of: *Ace* (5'-GCCAACCAGGGGATA-3'; 5'-CTGTCCGGTCATCTCTT-3'); *Kit* (5'-CTTGTGTCTTGGGAGAA-3'; 5'-GGTGCCATCCACTTAC-3'); *Prrm1* (5'-CGCAGC AAAAGCAGGAGCAG-3'; 5'-CATCGGACGGTGGCATTTTT-3'); and transcripts of *Hprt*, *Rsk3* (primer pair 145/146 (Fig. 1a, c); 5'-ATGGCCTGGGGATCATCTAC-3'; 5'-CACCGCTGCACACTGAGTA-3') and *Tcr* (primer pair 181/144 (Fig. 1a, c); 5'-TGGCATCTTATTGTCTAC-3'; 5'-TGCTCAAGCCAAAATCTGTG-3', spanning bp 1,441–2,263 of the cDNA 66bk/4-2). Primers were selected from different exons. RT-PCR was done according to standard procedures with 35 cycles of 30 s 94°C, 30 s 50°C and 30 s 72°C. *In situ* hybridization with digoxigenin-labelled RNA probes and staining were carried out as described³⁷ using the ORF of Tcr as the template for antisense RNA transcription. Paraffin sections were prepared from the testis of a *T/t*^{w5} male, hybridized and counterstained with eosin Y before mounting.

Transgenes

Tg4 consists of the testis-specific promoter of *Kit*, bp 45–683 (Styl)³⁸, sequences from bp 188 to bp 2,484 of 66bk/4-2 including the ORF of *Tcr*¹⁶, and an IRES-βgeo providing a polyadenylation signal. *Tg5* was constructed using the putative promoter of *Tcr* derived from t^{w12} , a 2.6-kb *KpnI/PmlI* fragment upstream of the splice donor of exon 1, ligated to bp 480–2,263 of the cDNA 66bk/4-2 encoding the ORF of *Tcr*¹⁶, followed by an IRES-hCD24 cassette and the polyadenylation signal of SV40-t including an intron. Transgene carriers were identified by PCR using the primer pair 309/310 (5'-CAGCCCATGAATC-CATC-3'; 5'-TGCCCTTCGGTCTGAAAG-3').

Protein kinase assay and western blot analysis

Fragments encoding the catalytic domains of Tcr¹⁶ and Smok⁺¹²⁹, and calmodulin III of mouse, were cloned into the vector pBAD/Thio-TOPO in frame with the V5-His tag and expressed in *Escherichia coli* according to the supplier's instructions (Invitrogen); proteins were purified under native conditions with Ni²⁺-NTA-Agarose (Qiagen). Kinase assays were performed in 20 mM HEPES, 2 mM EGTA, 5 mM MgCl₂, 100 mM NaCl, 1 mM β-mercaptoethanol with 0.2 mM ATP, 8 µCi ³²P-γ-ATP (3000 Ci per mmole) and 7.5 µg myelin basic protein as substrate, for 1 h at 30°C. Aliquots of the reactions were run on a 15% polyacrylamide gel, blotted on Hybond C extra membrane (Amersham) and exposed on a phosphorimager (FUJIX BAS 1000) to quantify the relative kinase activities. Calmodulin was the negative control. To measure the relative amounts of protein loaded, the filter was incubated first with an anti-V5 antibody (Invitrogen), then with anti-mouse IgG-POD, and developed with ECL western blotting detection reagents (Amersham) on a Lumi-Imager (Roch) to quantify the relative amounts of protein loaded.

Received 17 May; accepted 21 September 1999.

- Silver, L. M. The peculiar journey of a selfish chromosome: mouse *t* haplotypes and meiotic drive. *Trends Genet.* **9**, 250–254 (1993).
- Fraser, L. R. & Dudley, K. New insights into the *t*-complex and control of sperm function. *BioEssays* **21**, 304–312 (1999).
- Lyon, M. F. Transmission ratio distortion in mouse *t*-haplotypes is due to multiple *Distorter* genes acting on a *Responder* locus. *Cell* **37**, 621–628 (1984).
- Silver, L. M. & Remis, D. Five of nine genetically defined regions of mouse *t* haplotypes are involved in transmission ratio distortion. *Genet. Res. Camb.* **49**, 51–56 (1987).
- Lyon, M. F. Male sterility of the mouse *t*-complex is due to homozygosity of the *Distorter* genes. *Cell* **44**, 357–363 (1986).
- Olds-Clarke, P. & Johnson, L. *t* haplotypes in the mouse compromise sperm flagellar function. *Dev. Biol.* **155**, 14–25 (1993).
- Fox, H. S. *et al.* Molecular probes define different regions of the mouse *t* complex. *Cell* **40**, 63–69 (1985).
- Schimenti, J., Vold, L., Socolow, D. & Silver, L. M. An unstable family of large DNA elements in the center of the mouse *t* complex. *J. Mol. Biol.* **194**, 583–594 (1987).
- Rosen, L. L., Bullard, D. C., Silver, L. M. & Schimenti, J. C. Molecular cloning of the *t* complex *Responder* genetic locus. *Genomics* **8**, 134–140 (1990).
- Bullard, D. C., Ticknor, C. & Schimenti, J. C. Functional analysis of a *t* complex *Responder* locus transgene in mice. *Mamm. Genome* **3**, 579–587 (1992).

11. Schimenti, J. *et al.* A candidate gene family of the mouse *t* complex responder (*Tcr*) locus responsible for haploid effects on sperm function. *Cell* **55**, 71–78 (1988).
12. Ewulonu, U. K., Schimenti, K., Kuemerle, B., Magnuson, T. & Schimenti, J. Targeted mutagenesis of a candidate *t* complex responder gene in mouse *t* haplotypes does not eliminate transmission ratio distortion. *Genetics* **144**, 785–792 (1996).
13. Kispert, A., Stöger, R. J., Caparros, M. & Herrmann, B. G. The mouse *Rsk3* gene maps to the *Leh66* elements carrying the *t*-complex responder *Tcr*. *Mamm. Genome* **10**, 794–802 (1999).
14. Drewes, G., Ebneth, A., Preuss, U., Mandelkow, E. M. & Mandelkow, E. MARK, a novel family of protein kinases that phosphorylate microtubule-associated proteins and trigger microtubule disruption. *Cell* **89**, 297–308 (1997).
15. Hanks, S. K. & Hunter, T. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9**, 576–596 (1995).
16. Rugh, R. *The Mouse, Its Reproduction and Development* (Oxford Univ. Press, Oxford, 1990).
17. Peschon, J. J., Behringer, R. R., Brinster, R. L. & Palmiter, R. D. Spermatid-specific expression of protamine 1 in transgenic mice. *Proc. Natl Acad. Sci. USA* **84**, 5316–5319 (1987).
18. Langford, K. G. *et al.* Transgenic mice demonstrate a testis-specific promoter for angiotensin-converting enzyme. *J. Biol. Chem.* **266**, 15559–15562 (1991).
19. Albanesi, C. *et al.* A cell- and developmental stage-specific promoter drives the expression of a truncated *c-kit* protein during mouse spermatid elongation. *Development* **122**, 1291–1302 (1996).
20. Willison, K. & Ashworth, A. Mammalian spermatogenic gene expression. *Trends Genet.* **3**, 351–355 (1987).
21. Braun, R. E., Behringer, R. R., Peschon, J. J., Brinster, R. L. & Palmiter, R. D. Genetically haploid spermatids are phenotypically diploid. *Nature* **337**, 373–376 (1989).
22. Bennett, D., Alton, A. K. & Artzt, K. Genetic analysis of transmission ratio distortion by *t*-haplotypes in the mouse. *Genet. Res. Camb.* **41**, 29–45 (1983).
23. Lyon, M. F. Search for differences among *t* haplotypes in *distorter* and *responder* genes. *Genet. Res. Camb.* **55**, 13–19 (1990).
24. Tash, J. S., Kakar, S. S. & Means, A. R. Flagellar motility requires the cAMP-dependent phosphorylation of a heat-stable NP-40-soluble 56 kD protein, axonin. *Cell* **38**, 551–559 (1984).
25. Tash, J. S. *et al.* Identification, characterization, and functional correlation of calmodulin-dependent protein phosphatase in sperm. *J. Cell Biol.* **106**, 1625–1633 (1988).
26. Carrera, A. *et al.* Regulation of protein tyrosine phosphorylation in human sperm by a calcium/calmodulin-dependent mechanism: identification of A kinase anchor proteins as major substrates for tyrosine phosphorylation. *Dev. Biol.* **180**, 284–296 (1996).
27. Vijayaraghavan, S., Goueli, S. A., Davey, M. P. & Carr, D. W. Protein kinase A-anchoring inhibitor peptides arrest mammalian sperm motility. *J. Biol. Chem.* **272**, 4747–4752 (1997).
28. Bracho, G. E., Fritch, J. J. & Tash, J. S. Identification of flagellar proteins that initiate the activation of sperm motility in vivo. *Biochem. Biophys. Res. Com.* **242**, 231–237 (1998).
29. Cosson, J. A moving image of flagella: news and views on the mechanism involved in axonemal beating. *Cell Biol. Int.* **20**, 83–94 (1996).
30. Lader, E., Ha, H. S., O'Neill, M., Artzt, K. & Bennett, D. *tctex-1*: a candidate gene family for a mouse *t* complex sterility locus. *Cell* **58**, 969–979 (1989).
31. Huw, L. Y., Goldsborough, A. S., Willison, K. & Artzt, K. *Tctex2*: a sperm tail surface protein mapping to the *t*-complex. *Dev. Biol.* **170**, 183–194 (1995).
32. Harrison, A., Olds-Clarke, P. & King, S. M. Identification of the *t* complex-encoded cytoplasmic dynein light chain *tctex1* in inner arm II supports the involvement of flagellar dyneins in meiotic drive. *J. Cell. Biol.* **140**, 1137–1147 (1998).
33. Patel-King, R. S., Benashski, S. E., Harrison, A. & King, S. M. A *Clamydomonas* homologue of the putative murine *t* complex distorter *Tctex-2* is an outer arm dynein light chain. *J. Cell Biol.* **137**, 1081–1090 (1997).
34. Mazarakis, N. *et al.* Isolation and characterization of a testis-expressed developmentally regulated gene from the distal inversion of the mouse *t*-complex. *Development* **111**, 561–571 (1991).
35. Lyon, M. F. & Zenthor, J. Differences in or near the responder region of complete and partial mouse *t*-haplotypes. *Genet. Res. Camb.* **50**, 29–34 (1987).
36. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 1989).
37. Wilkinson, D. G. *In Situ Hybridization: A Practical Approach* (Oxford Univ. Press, Oxford, 1992).
38. Rossi, P. *et al.* A novel *c-kit* transcript, potentially encoding a truncated receptor, originates within a *kit* gene intron in mouse spermatids. *Dev. Biol.* **152**, 203–207 (1992).

Acknowledgements

We thank M. F. Lyon, K. Artzt, J.-L. Guenet and J. Nadeau for mouse strains and DNA; A.-M. Frischaut for a cosmid library of the genotype t^{w12}/t^{w12} ; B. Engst for technical assistance; S. Kuscher and H. Garbers for DNA injections; U. Stauffer for maintenance of strains; A. Bauer and M. Leitges for suggestions on kinase assays; D. Solter and R. Kemler for support; and S. Gasca, D. Solter, R. Cassada and M. Mallo for comments and suggestions on the manuscript.

Correspondence and requests for materials should be addressed to B.G.H. (e-mail: herrmann@immunbio.mpg.de). Gene sequences have been submitted to the EMBL Nucleotide Sequence Database; accession numbers: *Tcr*¹⁶, AJ245452; *Smok2*^{Balb}, AJ245453; *Smok2*^{tw5}, AJ245454; *Smok*^{tw12}, AJ245455; *Smok*^{tw129}, AJ245456.

Identification of *in vivo* substrates of the chaperonin GroEL

Walid A. Houry*, Dmitrij Frishman†, Christoph Eckerskorn‡, Friedrich Lottspeich§ & F. Ulrich Hartl*

* Department of Cellular Biochemistry, † GSF-Forschungszentrum für Umwelt und Gesundheit, Munich Information Center for Protein Sequences, and

§ Department of Protein Analytics, Max-Planck-Institut für Biochemie, Am Klopferspitz 18A, D-82152 Martinsried, Germany

‡ Toplab GmbH, Proteomics Division, Fraunhoferstrasse 18A, D-82152 Martinsried, Germany

The chaperonin GroEL has an essential role in mediating protein folding in the cytosol of *Escherichia coli*. Here we show that GroEL interacts strongly with a well-defined set of approximately 300 newly translated polypeptides, including essential components of the transcription/translation machinery and metabolic enzymes. About one third of these proteins are structurally unstable and repeatedly return to GroEL for conformational maintenance. GroEL substrates consist preferentially of two or more domains with $\alpha\beta$ -folds, which contain α -helices and buried β -sheets with extensive hydrophobic surfaces. These proteins are expected to fold slowly and be prone to aggregation. The hydrophobic binding regions of GroEL may be well adapted to interact with the non-native states of $\alpha\beta$ -domain proteins.

The three-dimensional fold of a protein is determined by the amino-acid sequence of the newly synthesized polypeptide chain. Although proteins can reach their folded states spontaneously, the efficiency of folding is often limited by the side reaction of aggregation. In the cell, misfolding and aggregation of proteins during their biogenesis and under conditions of cellular stress is prevented by molecular chaperones^{1–3}.

The chaperonin GroEL, along with its cofactor GroES, is the only chaperone system in *E. coli* that is essential under all growth conditions^{4,5}. GroEL is a homo-oligomer of 14 subunits, each of relative molecular mass 57,000 (M_r 57K), which are arranged into two heptameric rings, forming a cylindrical structure with two large cavities. Substrate protein, with hydrophobic amino-acid residues exposed, binds in the central cavity of the cylinder, engaging the hydrophobic surfaces exposed by the apical GroEL domains^{6,7}. The ring-shaped cofactor GroES then binds to the apical domains of GroEL in an ATP-dependent reaction, resulting in the displacement of the substrate into an enclosed cavity. Proteins up to M_r ~60K can fold in the GroEL–GroES cage in which aggregation is prevented. After ~10 s of folding, when the GroEL-bound ATP has been hydrolysed to ADP, ATP binding to the opposite ring of GroEL results in the dissociation of GroES and folded protein from GroEL. Proteins that are strongly dependent on GroEL may require several rounds of interaction with GroEL to reach their native state^{3,8,9}.

GroEL interacts *in vitro* with almost any non-native model protein⁷. However, *in vivo* GroEL is involved in the folding of only ~10% of newly translated polypeptides¹⁰, indicating a preference for a subset of *E. coli* proteins. We have identified a large number of the endogenous substrates of GroEL, many of which are structurally labile and interact with GroEL not only for the initial folding, but also for conformational maintenance during their lifetime in the cell, both under normal growth conditions and when the cell is exposed to heat stress. Structural analysis revealed that these proteins have a complex domain architecture, preferentially containing two or more domains with $\alpha\beta$ -folds.

A defined set of GroEL substrates

Pulse-chase labelling experiments were performed with live *E. coli* cells to analyse the set of newly synthesized proteins interacting with GroEL. At different times of chase, cells were lysed on ice in the presence of EDTA, which prevented the ATP-dependent release of protein substrates from GroEL, and GroEL–substrate complexes were isolated by immunoprecipitation with anti-GroEL antibodies¹⁰. Total soluble cytoplasmic proteins (Fig. 1a, b) and GroEL-bound

proteins (Fig. 1c, d) were separated on two-dimensional polyacrylamide gels (2D gels). Control experiments demonstrated the specificity of the anti-GroEL immunoprecipitations¹⁰ (see below).

At a rate of translation of 10–20 amino acids per s (ref. 11), a 15-s pulse with [³⁵S]-methionine allows for the labelling of an *E. coli* protein comprising about 150–300 amino acids. Proteins with M_r larger than ~40K, including GroEL itself, complete synthesis during the chase period with unlabelled methionine (Fig. 1b). In contrast to the expected complexity of newly labelled cytoplasmic proteins resolved on the 2D gels (Fig. 1a, b), the pattern of GroEL-bound proteins (referred to as GroEL substrates) was much simpler and surprisingly well defined (Fig. 1c, d). A core of no more than 250–300 proteins out of a total of ~2,500 cytoplasmic polypeptides were reproducibly observed in complex with GroEL immediately upon labelling (Fig. 1c), even after prolonged exposure. About half of these proteins were still detectable on GroEL after 10 min of chase (Fig. 1d), albeit in strongly reduced amounts relative to GroEL. Proteins that interact only very transiently with GroEL may have escaped detection in this analysis, but such proteins are not expected to be strongly dependent on the chaperonin for folding.

The pI distributions of total soluble cytoplasmic proteins and of GroEL substrates were found to be very similar (Fig. 1e). However, most GroEL substrates are larger than M_r ~20K (Fig. 1f), and so are likely to have several domains¹². The majority of substrates (79%) are smaller than M_r 60K (Fig. 1f).

Maintenance function of GroEL

A systematic analysis of the flux of newly labelled substrates through GroEL revealed three groups of proteins. About two-thirds of the proteins with M_r less than 60K (around 160 proteins) were released completely from GroEL during the chase with time constants between 20 s and 2 min (Fig. 2a), presumably reflecting the requirement of one to several rounds of GroEL binding and release to reach their native state. In contrast, for a core group of around 100 proteins of M_r < 60K, a fraction of the population of a particular protein (5–30% of the initial amount recovered) remained associated with GroEL after the chase (Fig. 2b). In addition, several proteins larger than 60K were also observed to bind GroEL, but these were inefficiently released from the chaperonin (Fig. 2c). These proteins exceed the size limitation of the GroEL–GroES cage and cannot be enclosed in the GroEL cavity by GroES¹³. However, these proteins contribute only 6% of the total mass of proteins interacting with GroEL, as judged by the analysis of Coomassie-stained 2D gels.

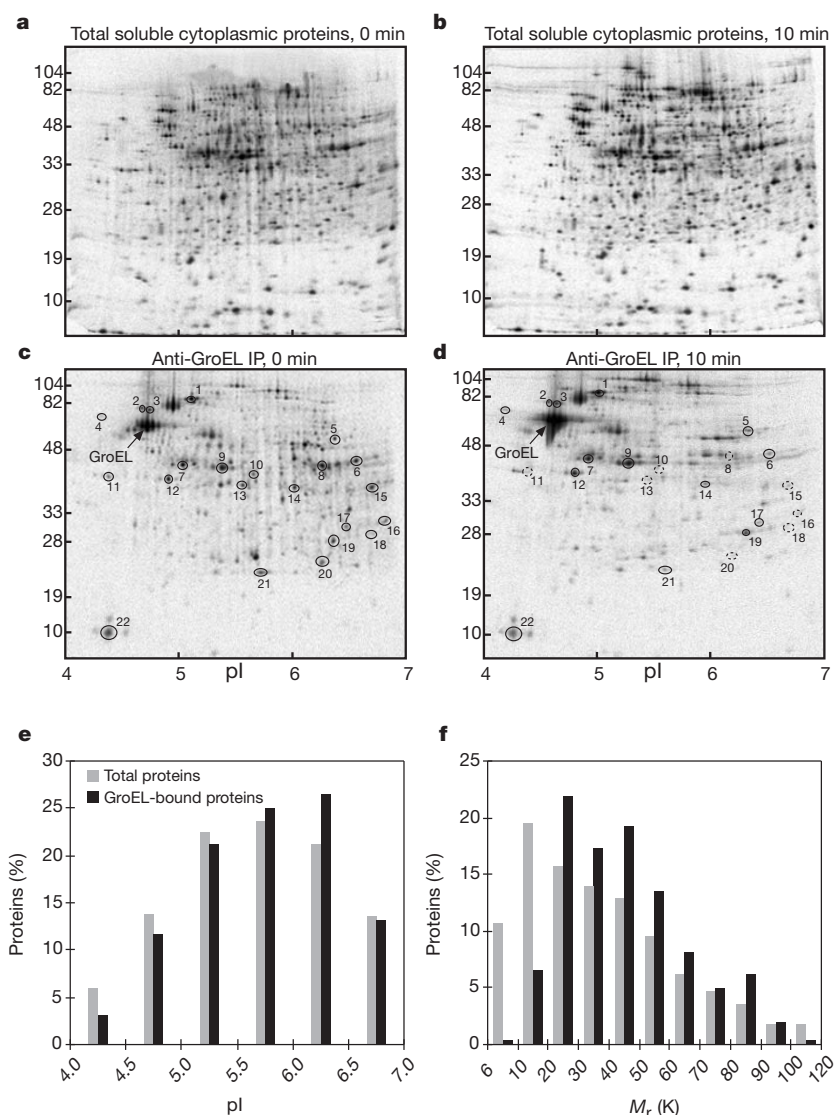


Figure 1 2D-gel analysis of newly translated proteins that transit through GroEL. **a–d.** *E. coli* MG1655 cells grown at 30 °C to mid-log phase were pulsed for 15 s and chased for 0–10 min. Equal amounts of total soluble cytoplasmic proteins (**a, b**) or immunoprecipitated proteins (**c, d**) were separated on 2D gels. In **c, d**, circles refer to proteins whose properties of interaction with GroEL are discussed in Figs 2 and 3. **e, f.** pI

(**e**) and M_r (**f**) distributions of total soluble cytoplasmic proteins in *E. coli* (grey bars; mean pI, 5.7; mean M_r , 37.5K) and of GroEL substrates (black bars; mean pI, 5.8; mean M_r , 45.0K) at 0 min of chase (**a, c**). Similar distributions were obtained at the other time points.

The persistence on GroEL of a fraction of newly synthesized polypeptides indicates that there are pre-existing proteins in *E. coli* (as opposed to newly synthesized proteins) that may undergo repeated cycles of GroEL binding and release. Indeed, when cells were incubated with [35 S]-methionine for 5 min or longer and then chased with unlabelled methionine, the amount and pattern of labelled polypeptides bound to GroEL was very similar, from 10 min of chase up to more than 2 hours, provided that the increase in cell mass during the chase time was corrected for (the doubling time of cells is 1 h) (Fig. 3a, left). These proteins were specifically bound to GroEL: they were not recovered with the chaperonin when the cell extract was incubated with SDS¹⁰, and they were released from GroEL when ATP and GroES were added (Fig. 3a, right). Rebinding to GroEL was not observed in the dilute cell lysates, confirming the earlier conclusion that the interactions occurred in the intact cells, rather than during lysis¹⁰. In contrast to the pre-existing GroEL substrates of $M_r < 60$ K, only a fraction of the GroEL-bound proteins of $M_r > 60$ K were released from the chaperonin when ATP and GroES were added to the lysate (Fig. 3a, right). These proteins may represent dead-end species.

The 2D gel pattern of pre-existing GroEL substrates (Fig. 3b) was very similar to that of newly translated proteins recovered in complex with GroEL after long chase times (Fig. 1d), indicating that the pre-existing proteins may also be using the chaperonin for initial folding. From the quantification of GroEL immunoprecipitates, about 1% of the cellular content of these proteins is associated with GroEL at any one time, assuming average abundance. Given a GroEL folding reaction of 10 s, the entire population of a pre-existing GroEL substrate will thus cycle on and off GroEL about four times per doubling time of the cells (1 h). The normal concentration of the GroEL oligomer in the cytosol is $\sim 3 \mu\text{M}$ ¹⁴, but each ring of GroEL may be active in polypeptide binding¹⁵. At a total protein concentration of 200 g l^{-1} in the cytosol¹⁶ and an average M_r of cytosolic proteins of 33K, the set of pre-existing GroEL substrates would occupy about 30% of the total polypeptide-binding capacity of GroEL.

It seemed plausible that the pre-existing proteins interacting with GroEL are relatively unstable, populating partly folded states to a significant extent *in vivo*. This idea could be confirmed by analysing the interaction of pre-existing proteins with GroEL under condi-

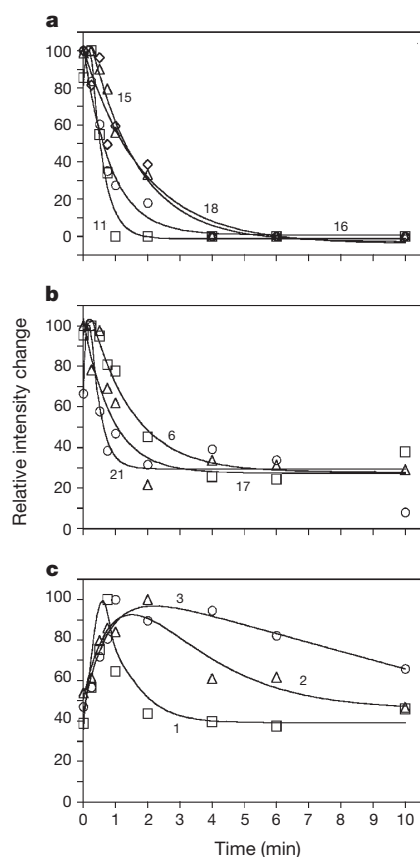


Figure 2 Kinetics and extent of release from GroEL of newly translated proteins. Kinetics of flux through GroEL of some newly translated proteins (indicated by circles in Fig. 1c, d) are shown. For a given kinetic trace, the maximum value obtained was arbitrarily set at 100. Solid lines are traces through the experimental points based on single- or double-exponential fits to the data. **a**, **b**, Examples of proteins of $M_r < 60K$ that are either completely (**a**) or partly (**b**) released from GroEL, with time constants of decay between 20 s and 2 min. **c**, Examples of proteins of $M_r > 60K$ that exhibit very slow decay kinetics (time constant > 2 min) with inefficient release from GroEL. For clarity, the flux kinetics of only 10 representative proteins of Fig. 1c, d are shown.

tions of heat stress (Fig. 3c, d). For this experiment, spheroplasts were used rather than cells, as they can be lysed rapidly in a hypo-osmotic buffer. It was found that, under heat stress at 43 °C, predominantly the same set of pre-existing proteins interacted more extensively with GroEL (compare Fig. 3c and d), although GroEL was about threefold more abundant at 43 °C than at 30 °C. The total amount of pre-existing substrates on GroEL increased roughly 5-fold, varying between 3- and 12-fold for individual polypeptides (data not shown). Thus, under heat stress, the fraction of GroEL devoted to conformational maintenance of proteins approximately doubles.

GroEL substrates

To identify the proteins that interact with GroEL for initial folding and conformational maintenance, large-scale immunoprecipitations of GroEL-substrate complexes were performed from *E. coli* cells under the conditions described above. Isolated complexes were separated on 2D gels (Fig. 4) and protein spots were analysed by mass spectroscopy. This procedure identified unequivocally 52 of the most abundant GroEL-bound proteins (Table 1), including several components of the transcription/translation machinery, such as subunits of the RNA polymerase, elongation factor Tu, and several aminoacyl transfer RNA synthetases, as well as a variety of important metabolic enzymes. We verified that these proteins transit GroEL on synthesis by 2D-gel analysis of a large-scale

immunoprecipitation of GroEL complexes that had been mixed with an immunoprecipitate from pulse-labelled cells prepared as in Fig. 1c (data not shown). The extent to which these proteins are dependent on GroEL and GroES for their folding remains to be determined. A preliminary analysis of the flux kinetics through GroEL indicates that the proteins NUSA (M_r 55K), S-adenosyl-methionine synthase (42K), elongation factor Tu (43K), RNA polymerase α -chain (37K) and 50S ribosomal protein L7/L12 (12K) interact with GroEL for initial folding and return to GroEL for conformational maintenance. They are released from GroEL *in vitro* when GroES and ATP are added (data not shown).

We observed that the α and β subunits of the RNA polymerase core complex and the ω subunit of the holoenzyme¹⁷ are substrates of GroEL (Table 1), in accord with previous genetic and biochemical evidence^{18,19}. The α subunit of the RNA polymerase represents 0.6% of the total cellular protein²⁰. Under normal growth conditions, ~2% of the subunit was found to be associated with the chaperonin at any one time (Fig. 4), in agreement with GroEL being involved in its maintenance.

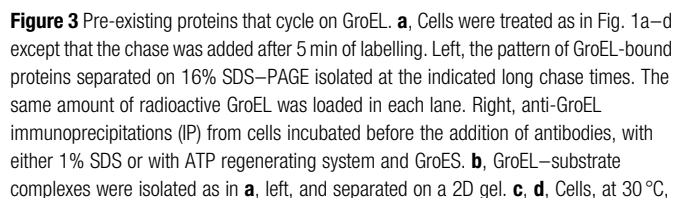
A preferred structural motif

Sequence analysis of the 52 identified GroEL substrates failed to reveal statistically significant consensus sequences or clusters of consensus sequences. Given the preference of GroEL for a subset of *E. coli* proteins *in vivo*, we considered the possibility that many of these proteins contain common structural motifs or folds that form the basis of their interaction with the chaperonin.

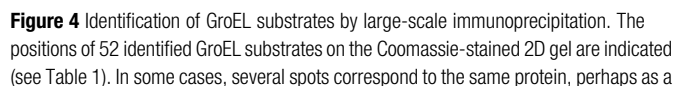
To investigate this possibility, we focused the analysis on the GroEL substrates with known 3D structures or with homologues of known structures, using the protein domain-classification databases SCOP²¹ and CATH²². For every protein, domains were classified independently through sequence homology to domains in these databases. Proteins containing a stretch of more than 100 contiguous amino acids not covered by sequence homology to domains in SCOP (or CATH) were excluded from the analysis. Of the 52 identified GroEL substrates, 24 were amenable to structural analysis (18 proteins using homologies with $> 17\%$ identities, or 24 with $> 13\%$ identities; see Methods and Table 1). The 24 proteins were representative GroEL substrates, based on their M_r distributions and predictions of secondary structure. Both SCOP and CATH gave similar results. We found that, with high statistical significance, GroEL substrates preferentially contain several $\alpha\beta$ domains compared with *E. coli* proteins (Fig. 5a). Of the multidomain GroEL substrates, 13 of 17 have at least two $\alpha\beta$ domains (Table 1). In contrast, no significant preference for all α , all β , discontinuous domains, or oligomeric state was found for GroEL substrates, compared with *E. coli* proteins, although these types of proteins are not excluded from the set of substrates.

The preference of GroEL for multiple $\alpha\beta$ domains in the GroEL substrates, compared with soluble *E. coli* proteins, is not due to a difference in M_r distribution between the two sets. If the structural analysis is restricted to proteins of $M_r > 20K$, to match the size bias of GroEL-bound proteins (Fig. 1f), the same significant preference for multiple $\alpha\beta$ domains is obtained for the GroEL substrates. Furthermore, based on their M_r distribution and predicted secondary structure, the set of *E. coli* proteins from SCOP or CATH faithfully represent the set of total soluble proteins in terms of their fold classification (data not shown).

The most common domain architectures in GroEL substrates were those of the three-layer $\alpha\beta\alpha$ sandwich and, with higher preference, the two-layer $\alpha\beta$ sandwich (data not shown). As a common occurrence in such structural motifs²³, the β -sheets expose a hydrophobic surface packed against the hydrophobic surfaces of α -helices (Fig. 5b). These β -sheets and the corresponding hydrophobic faces of the α -helices would provide ideal hydrophobic surfaces to mediate high-affinity interactions with the apical domains of GroEL. Several stringent model substrates of GroEL



GroEL IP



150 © 1999 Macmillan Magazines Ltd NATURE | VOL 402 | 11 NOVEMBER 1999 | www.nature.com

that are typically used for *in vitro* studies, including ornithine transcarbamylase²⁴, malate dehydrogenase²⁵, rhodanese¹⁰ and RUBISCO²⁶, belong to this category of $\alpha\beta$ proteins.

Substrate interaction with GroEL

The finding of a preferred domain topology in GroEL substrates provides insight into why and how these proteins interact with the chaperonin. For $\alpha\beta$ domains, the formation of the β -sheet is expected to be the most difficult step in the folding process because, unlike the formation of an α -helix, assembly of the β -sheet requires the formation of a large number of specific long-range contacts in the proper orientation. In general, these domains are expected to exhibit relatively slow folding rates²⁷, and misfolding or kinetic trapping may occur either through the improper packing of helices and sheets within one domain or between domains, or owing to the packing of helices in one molecule against a sheet in another molecule. As GroEL substrates consist preferentially of two or more $\alpha\beta$ domains, these proteins may be particularly prone to aggregation as a consequence of 3D domain swapping²⁸.

The estimated number of proteins with multiple $\alpha\beta$ domains in

the *E. coli* cytoplasm is between 200 and 600. GroEL may assist in the initial folding and conformational maintenance of a subset of these proteins in several ways. The hydrophobic residues exposed on two flexible helices and a loop region in the apical domain of GroEL, which mediate polypeptide binding⁶, could provide an adjustable scaffold to stabilize the β -sheet of the substrate protein, essentially by acting as a substitute for the helices in the native protein. Subsequently, folding of a single protein molecule would proceed upon GroES-mediated displacement into the enclosed GroEL–GroES cavity. If helices and sheet(s) in a substrate protein have been packed improperly, multiple apical domains of GroEL could interfere by binding the helices²⁹, the β -sheet or both, thereby dissociating their improper interaction. Similarly, for structurally labile $\alpha\beta$ proteins, partial unfolding would lead to the exposure of the hydrophobic surfaces of the helices and the buried β -sheets which can be bound by GroEL. These proteins would therefore continually require GroEL for conformational maintenance and refolding. For several model proteins, the native state has been shown to be in equilibrium with unfolding intermediates that bind to GroEL^{30–32}.

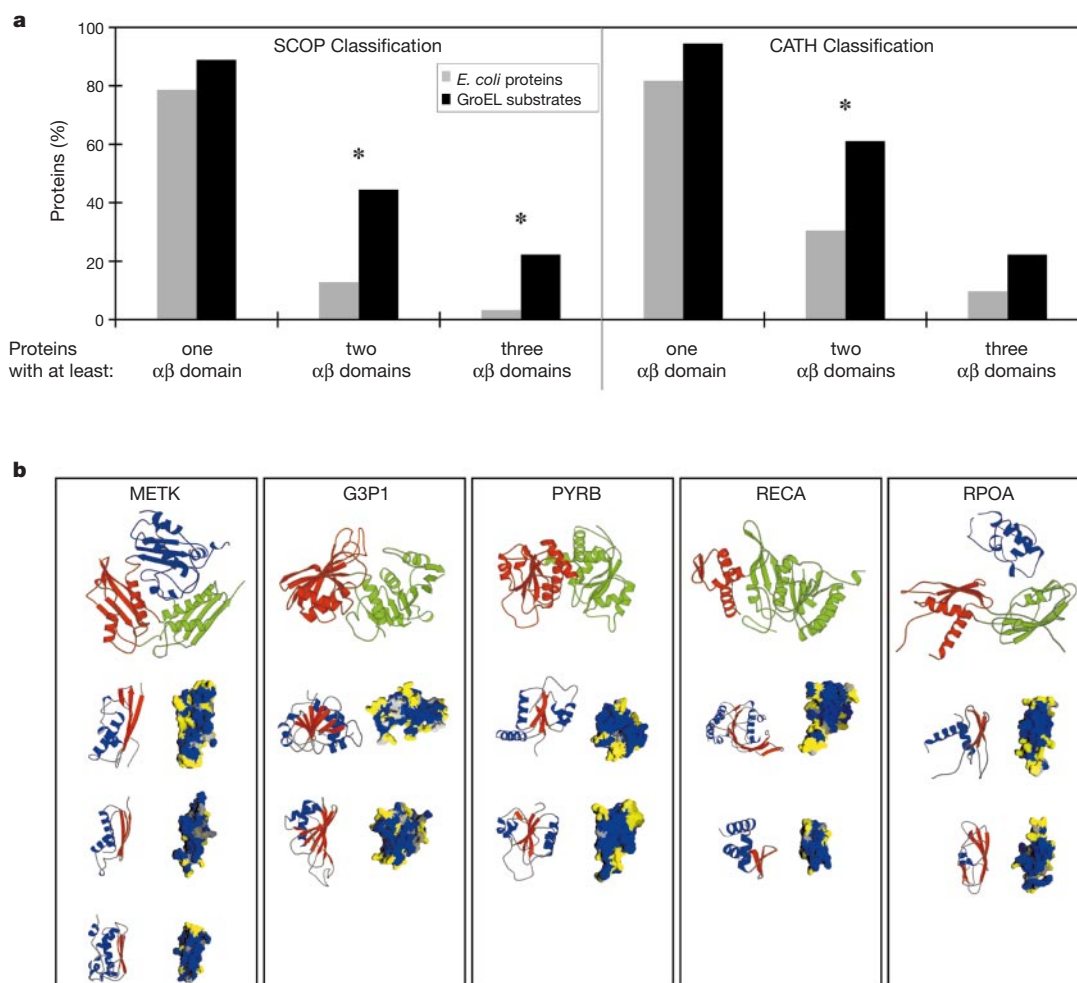


Figure 5 Structural classification of GroEL substrates. **a**, SCOP or CATH domain-classification databases were used to determine the number of proteins in *E. coli* (grey bars, ~400 proteins) or in the set of identified GroEL substrates (black bars, 18 proteins) with at least one, two or three $\alpha\beta$ domains. (A very similar result was obtained when including the additional six GroEL substrates which have a lower homology score.) The differences observed (asterisk) between the *E. coli* set and the GroEL substrates were found to be statistically significant at the 95% confidence limit by using equations developed for sampling statistics (see ref. 47). **b**, Structures of representative GroEL substrates. Top, the whole protein subunit chain, colour coded by domains. Domains that

are classified as three-layer $\alpha\beta\alpha$ or two-layer $\alpha\beta$ sandwiches according to CATH, except for RNA polymerase α chain (RPOA) whose structural classification is based on ref. 48 and was not included in the statistical analysis, are depicted individually with helices in blue, strands in red and other secondary structure elements in grey. The hydrophobic face of the buried sheet for each of the domains is shown as molecular surface colour-coded according to hydrophobicity (yellow is hydrophilic and blue is hydrophobic) based on a scale defined in GRASP⁴⁹. The names of GroEL substrates are abbreviated according to Table 1. Protein Database files used are METK, 1mxb; G3P1, 1gadQ; PYRB, 1raiA; RECA, 2reb; RPOA, 1bdfA + 1coo.

Table 1 Substrates of GroEL

	SwissProt ID	<i>M_r</i> (k)	Domain structure
Amino-acid metabolism			
Chorismate biosynthesis			
Phospho-2-dehydro-3-deoxyheptonate aldolase, Phe-sensitive	AROG_ECOLI	38.0	
Phospho-2-dehydro-3-deoxyheptonate aldolase, Trp-sensitive	AROH_ECOLI	38.7	
Methionine biosynthesis and conversion			
Tetrahydropteroyltryglutamate methyltransferase	METE_ECOLI	84.5	
5,10-Methylenetetrahydrofolate reductase	METF_ECOLI	33.1	
S-adenosylmethionine synthetase	METK_ECOLI	42.0	3αβ
Val, leu, Ile biosynthesis			
Acetolactate synthase isozyme III large subunit	ILVI_ECOLI	63.0	3αβ + 1f
2-Isopropylmalate synthase	LEU1_ECOLI	57.2	
Amino-acid catabolism and cleavage			
D-amino acid dehydrogenase small subunit	DADA_ECOLI	47.6	2αβ
Glutamate decarboxylase-α	DCEA_ECOLI	52.7	
Dihydrolipoamide dehydrogenase	DLDH_ECOLI	50.6	3αβ
Threonine 3-dehydrogenase	TDH_ECOLI	37.2	2αβ
Sugar metabolism			
Galactitol-tagatose pathway			
Galactitol-1-phosphate 5-dehydrogenase	GATD_ECOLI	37.4	2αβ
Tagatose-6-phosphate kinase gatZ	GATZ_ECOLI	47.1	
Tagatose-bisphosphate aldolase gatY	GATY_ECOLI	31.1	1αβ
Glycolysis pathway			
Enolase	ENO_ECOLI	45.5	2αβ
Glyceraldehyde 3-phosphate dehydrogenase A	G3P1_ECOLI	35.4	2αβ
N-acetylglucosamine utilization			
NAGD protein	NAGD_ECOLI	27.2	1αβ + 1α
Pyruvate oxidation			
Phosphate acetyltransferase	PTA_ECOLI	77.0	
O-antigen biosynthesis			
UDP-galactopyranose mutase	GLF_ECOLI	43.0	
Capsular synthesis			
Capsular-synthesis regulator component B	RCSB_ECOLI	23.7	1αβ
Other metabolism			
Purine biosynthesis			
Phosphoribosylaminoimidazole-succinocarboxamide synthase	PUR7_ECOLI	27.0	
Pyrimidine biosynthesis and salvage			
Aspartate carbamoyltransferase catalytic chain	PYRB_ECOLI	34.3	2αβ
Uracil phosphoribosyltransferase	UPP_ECOLI	22.5	1αβ
Thiamin biosynthesis			
Phosphomethylpyrimidine kinase	THID_ECOLI	28.6	
THIG protein	THIG_ECOLI	29.7	1αβ
THIH protein	THIH_ECOLI	43.3	
Biotin biosynthesis			
8-Amino-7-oxononanoate synthase	BIOF_ECOLI	41.6	
Lipoate biosynthesis			
Lipoic acid synthetase	LIPA_ECOLI	36.1	
Transcription and translation			
RNA polymerase			
DNA-directed RNA polymerase α chain	RPOA_ECOLI	36.5	2αβ + 1α
DNA-directed RNA polymerase β chain fragment	RPOB_ECOLI	150.6	
DNA-directed RNA polymerase ω chain	RPOZ_ECOLI	10.2	
Ribosomal proteins			
50S ribosomal protein L7/L12	RL7_ECOLI	12.2	1αβ
50S ribosomal protein L9	RL9_ECOLI	15.8	2αβ
30S ribosomal protein S2	RS2_ECOLI	26.6	
tRNA synthetases			
Phenylalanyl-tRNA synthetase α chain	SYFA_ECOLI	36.8	1αβ
Phenylalanyl-tRNA synthetase β chain	SYFB_ECOLI	87.4	3αβ + 1β
Glycyl-tRNA synthetase β chain	SYGB_ECOLI	76.7	
Threonyl-tRNA synthetase	SYT_ECOLI	74.0	
Transcription termination and antitermination			
NUSA protein	NUSA_ECOLI	54.9	
NUSG protein	NUSG_ECOLI	20.5	
Translation			
Elongation factor Tu	EFTU_ECOLI	43.2	1αβ + 2β

continued

Table 1 continued

Miscellaneous	SwissProt ID	M_r (K)	Domain structure
Stress response proteins			
ATP-dependent Clp protease ATP-binding subunit ClpX	CLPX_ECOLI	46.2	1 $\alpha\beta$ + 1 β
Heat-shock protein GrpE	GRPE_ECOLI	21.8	
Stringent starvation protein B	SSPB_ECOLI	18.3	
DNA manipulation			
DNA gyrase subunit A	GYRA_ECOLI	97.0	2 $\alpha\beta$
RecA protein	RECA_ECOLI	37.8	
Cell-division related			
Cell division inhibitor minD	MIND_ECOLI	29.5	1 $\alpha\beta$
Intake of inorganic phosphate			
Phosphate transport ATP-binding protein PSTB	PSTB_ECOLI	29.0	
Iron storage			
Ferritin	FTN_ECOLI	19.4	1 α
Function not established			
Protein YCFV	YCFV_ECOLI	24.9	1 $\alpha\beta$ + 1 α
Protein in XERC-UVRD intergenic region	YIGB_ECOLI	27.1	
Protein in OXMY-DEOC intergenic region	YJJU_ECOLI	39.8	

The 52 identified GroEL substrates are grouped into different categories according to function, based on the information available in the SwissProt database⁴⁵, EcoCyc database⁴⁶ and the general literature. For each substrate protein, the following information is listed: name, which is generally the same as that used in SwissProt; SwissProt ID; M_r (K); the number and structural classification of its domains (based on CATH, except for RPOA; see Methods and legend of Fig. 5). $\alpha\beta$ collectively refers to the $\alpha\beta$ class of domains. α , β , and Γ refer to mainly α , mainly β , and a few secondary structure domains, respectively.

The typical GroEL substrate

The role of GroEL in assisting the folding or maintenance of a particular protein *in vivo* is expected to depend on the presence of appropriate binding sites for GroEL in the folding or unfolding intermediates of that protein, and on the rate at which these sites are buried during folding or refolding, respectively. Our findings indicate that a typical GroEL substrate has an M_r of between 20K and 60K, and consists preferentially of several $\alpha\beta$ domains with buried hydrophobic β -sheets. The non-native states of these topologically complex proteins are likely to expose extensive hydrophobic surfaces that could be recognized by GroEL, either during folding or on misfolding. Co-expression of chaperonin may provide a rational strategy to improve the folding efficiency of foreign proteins expressed in *E. coli* when these proteins contain multiple $\alpha\beta$ domains. □

Methods

Cell labelling, immunoprecipitation and 2D-gel analysis

Pulse-chase experiments on *E. coli* MG1655 cells (F λ) followed by immunoprecipitation of GroEL–substrate complexes were carried out as described¹⁰. Radioactive proteins on SDS–PAGE or on 2D gels were visualized by using a FUJIFILM FLA-2000 phosphor-imager. Control immunoprecipitations were performed from cell lysates that were heated to 95 °C for ~1 min with 1% SDS and then diluted 10 times into lysis buffer containing 0.5% Triton X-100 (Fig. 3a, right, first lane). For controls of ATP with GroES (Fig. 3a, right, second lane), cells were lysed hypo-osmotically without EDTA, and then 500 μ l of the cytoplasmic fraction was incubated for 30 min at room temperature with 10 mM ATP, 200 μ g ml⁻¹ creatine kinase and 10 mM creatine phosphate with 0.2 μ M purified GroES³⁵. Samples for 2D-gel analysis were focused using a 13-cm Immobiline DryStrip pH 4–7 L and the Multiphor II system (Pharmacia), followed by SDS–PAGE^{34,35}. Kinetics of flux through GroEL of newly translated proteins (Fig. 2) were determined by measuring the volume of the radioactive spots corresponding to each protein from the 2D gel of the anti-GroEL immunoprecipitation at different chase times using ImageMaster 2D Elite (Pharmacia). The volume obtained at each time point was normalized to the GroEL volume at that time point and scaled according to the increase in the intensity of the GroEL band as a function of chase time visualized on SDS–PAGE.

Identification of GroEL substrates

GroEL–substrate complexes were isolated by large-scale immunoprecipitation from 2-l cultures of MG1655 cells grown in minimal media to mid-log phase. Cells were lysed hypo-osmotically on ice and about 16 ng of affinity-purified goat anti-GroEL antibodies crosslinked to protein-G-Sepharose were used in the immunoprecipitations. Proteins were eluted from the beads using 8 M urea/4% CHAPS. Portions of the cytoplasmic fraction taken before and after immunoprecipitation of GroEL complexes were exchanged into 8 M urea/4% CHAPS. Samples were concentrated to about 300 μ l for 2D-gel analysis (Amicon concentrators with M_r 3K cutoff) and loaded into 18-cm Immobiline DryStrip pH 3–10 NL for isoelectric focusing, followed by 9–16% SDS–PAGE. The pattern of spots

on the 2D gel from the large-scale immunoprecipitation was reproducibly obtained in three separate repeats. Approximately 110 protein spots in the 2D gel were treated with trypsin followed by peptide-mass fingerprint analysis using MALDI-TOF mass spectrometry³⁶. All proteins identified were found to be *E. coli* proteins, which further supported their correct identification.

Sequence and structure analysis

Pairwise and multiple sequence alignments were carried out using PSI-BLAST³⁷ (BLOSUM62 matrix, filtering using SEG³⁸, default gap penalty, and an e-value of 0.01). An all-against-all sequence comparison of the identified GroEL substrates gave only three pairs of proteins with significant sequence similarity: GATD_ECOLI and TDH_ECOLI; AROG_ECOLI and AROH_ECOLI, and PSTB_ECOLI and YCFV_ECOLI. Because up to one third of the *E. coli* proteins are estimated to have resulted from gene duplication events³⁹, GroEL displays no preference for a particular gene family.

For secondary structure predictions, membrane proteins, defined as having two or more predicted transmembrane regions, were first identified using ALOM⁴⁰. Structural predictions of the remaining globular proteins were then calculated using PREDATOR⁴¹. Proteins were attributed to $\alpha\beta$, all α , all β , or irregular category classes⁴².

For homology-based fold assignments, iterative similarity searches using PSI-BLAST were carried out with each query protein sequence against SCOP²¹ or CATH²². Duplicate hits from different parts of discontinuous domains were removed. For *E. coli*, 492 proteins were obtained for the structural analysis using SCOP and 401 using CATH; for the group of 52 identified GroEL substrates, this number was 18 using either database. Homologies obtained with this procedure had identities of 17–100%. By using PSI-BLAST searches against a combination of the full-protein sequence database and SCOP or CATH domains⁴³, six additional proteins were structurally assigned from the set of GroEL substrates through homologies with identities between 13 and 17%.

In CATH, protein domains are divided into four classes of $\alpha\beta$, mainly α , mainly β , and few secondary structures. Each class is then subdivided into different architectures. The most common architectures of the $\alpha\beta$ class are those of the three-layer $\alpha\beta\alpha$ sandwich, consisting of a relatively flat β -sheet sandwiched between two layers of α helices, and the two-layer $\alpha\beta$ sandwich, consisting of a relatively flat β -sheet packed against a layer of α helices. In SCOP, as in CATH, a similar but more specific division of protein domains into different classes is found. The main domain classes found in SCOP are: α and β (typically consisting of a central β -sheet with repeated units of β strand– α helix– β strand), α plus β (containing segregated α and β regions), all α , all β , and multidomain classes. When using SCOP, we grouped the (α and β) and (α plus β) classes into one $\alpha\beta$ class to be consistent with the CATH classification. Results of the structural and functional characterization of the 52 identified GroEL substrates will be available on the PEDANT server⁴⁴ (<http://pedant.mips.biochem.mpg.de>).

Received 18 May; accepted 22 September 1999.

- Georgopoulos, C. & Welch, W. J. Role of the major heat shock proteins as molecular chaperones. *Annu. Rev. Cell Biol.* **9**, 601–634 (1993).
- Ellis, R. J. Roles of molecular chaperones in protein folding. *Curr. Opin. Struct. Biol.* **4**, 117–122 (1994).
- Hartl, F. U. Molecular chaperones in cellular protein folding. *Nature* **381**, 571–580 (1996).
- Fayet, O., Ziegelhoffer, T. & Georgopoulos, C. The GroES and GroEL heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. *J. Bacteriol.* **171**, 1379–1385 (1989).

5. Horwich, A. L., Low, K. B., Fenton, W. A., Hirshfield, I. N. & Furtak, K. Folding *in vivo* of bacterial cytoplasmic proteins: role of GroEL. *Cell* **74**, 909–917 (1993).
6. Fenton, W. A., Kashi, Y., Furtak, K. & Horwich, A. L. Residues in chaperonin GroEL required for polypeptide binding and release. *Nature* **371**, 614–619 (1994).
7. Coyle, J. E., Jaeger, J., Gross, M., Robinson, C. V. & Radford, S. E. Structural and mechanistic consequences of polypeptide binding by GroEL. *Fold. Design* **2**, R93–R104 (1997).
8. Sigler, P. B. *et al.* Structure and function in GroEL-mediated protein folding. *Annu. Rev. Biochem.* **67**, 581–608 (1998).
9. Ranson, N. A., White, H. E. & Saibil, H. R. Chaperonins. *Biochem. J.* **333**, 233–242 (1998).
10. Ewalt, K. L., Hendrick, J. P., Houry, W. A. & Hartl, F. U. *In vivo* observation of polypeptide flux through the bacterial chaperonin system. *Cell* **90**, 491–500 (1997).
11. Pedersen, S. *Escherichia coli* ribosomes translate *in vivo* with variable rate. *EMBO J.* **3**, 2895–2898 (1984).
12. Xu, D. & Nussinov, R. Favorable domain size in proteins. *Fold. Design* **3**, 11–17 (1998).
13. Xu, Z. H., Horwich, A. L. & Sigler, P. B. The crystal structure of the asymmetric GroEL–GroES–(ADP) chaperonin complex. *Nature* **388**, 741–750 (1997).
14. Ellis, R. J. & Hartl, F. U. Protein folding in the cell: Competing models of chaperonin function. *FASEB J.* **10**, 20–26 (1996).
15. Rye, H. S. *et al.* GroEL–GroES cycling: ATP and nonnative polypeptide direct alternation of folding-active rings. *Cell* **97**, 325–338 (1999).
16. Ellis, R. J. Molecular chaperones: avoiding the crowd. *Curr. Biol.* **7**, R531–R533 (1997).
17. Gentry, D. R. & Burgess, R. R. The cloning and sequence of the gene encoding the omega subunit of *Escherichia coli* RNA polymerase. *Gene* **48**, 33–40 (1986).
18. Wada, M., Fujita, H. & Itikawa, H. Genetic suppression of a temperature-sensitive groES mutation by an altered subunit of RNA polymerase of *Escherichia coli* K-12. *J. Bacteriol.* **169**, 1102–1106 (1987).
19. Ziemienowicz, A. *et al.* Both the *Escherichia coli* chaperone systems, GroEL/GroES and DnaK/DnaJ/GrpE, can reactivate heat-treated RNA polymerase. Different mechanisms for the same activity. *J. Biol. Chem.* **268**, 25425–25431 (1993).
20. VanBogelen, R. A., Sankar, P., Clark, R. L., Bogan, J. A. & Neidhardt, F. C. The gene–protein database of *Escherichia coli*: edition 5. *Electrophoresis* **13**, 1014–1054 (1992).
21. Hubbard, T. J. P., Ailey, B., Brenner, S. E., Murzin, A. G. & Chothia, C. SCOP: a structural classification of proteins database. *Nucleic Acids Res.* **27**, 254–256 (1999).
22. Orengo, C. A. *et al.* The CATH database provides insights into protein structure/function relationships. *Nucleic Acids Res.* **27**, 275–279 (1999).
23. Eisenberg, D., Wilcox, W. & McLachlan, A. D. Hydrophobicity and amphiphilicity in protein structure. *J. Cell. Biochem.* **31**, 11–17 (1986).
24. Cheng, M. Y. *et al.* Mitochondrial heat-shock protein hsp60 is essential for assembly of proteins imported into yeast mitochondria. *Nature* **337**, 620–625 (1989).
25. Rye, H. S. *et al.* Distinct actions of *cis* and *trans* ATP within the double ring of the chaperonin GroEL. *Nature* **388**, 792–798 (1997).
26. Viitanen, P. V. *et al.* Chaperonin-facilitated refolding of the ribulosebiphosphate carboxylase and ATP hydrolysis by chaperonin 60 (GroEL) are K⁺ dependent. *Biochemistry* **29**, 5665–5671 (1990).
27. Plaxco, K. W., Simons, K. T. & Baker, D. Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* **277**, 985–994 (1998).
28. Schlunegger, M. P., Bennett, M. J. & Eisenberg, D. Oligomer formation by 3D domain swapping: a model for protein assembly and misassembly. *Adv. Protein Chem.* **50**, 61–122 (1997).
29. Landry, S. J., Jordan, R., McMacken, R. & Gierasch, L. M. Different conformations for the same polypeptide bound to chaperones DnaK and GroEL. *Nature* **355**, 455–457 (1992).
30. Lamiet, A. A., Ziegelhoffer, T., Georgopoulos, C. & Pluckthun, A. The *Escherichia coli* heat shock proteins GroEL and GroES modulate the folding of the beta-lactamase precursor. *EMBO J.* **9**, 2315–2319 (1990).
31. Vitanen, P. V., Donaldson, G. K., Lorimer, G. H., Lubben, T. H. & Gatenby, A. A. Complex interactions between the chaperonin 60 molecular chaperone and dihydrofolate reductase. *Biochemistry* **30**, 9716–9723 (1991).
32. Smith, K. E., Voziyan, P. A. & Fisher, M. T. Partitioning of rhodanese onto GroEL–chaperonin binds a reversibly oxidized form derived from the native protein. *J. Biol. Chem.* **273**, 28677–28681 (1998).
33. Hayer-Hartl, M. K., Weber, F. & Hartl, F. U. Mechanism of chaperonin action: GroES binding and release can drive GroEL-mediated protein folding in the absence of ATP hydrolysis. *EMBO J.* **15**, 6111–6121 (1996).
34. Bjellqvist, B., Pasquali, C., Ravier, F., Sanchez, J. C. & Hochstrasser, D. A nonlinear wide-range immobilized pH gradient for two-dimensional electrophoresis and its definition in a relevant pH scale. *Electrophoresis* **14**, 1357–1365 (1993).
35. Gorg, A., Postel, W. & Gunther, S. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* **9**, 531–546 (1988).
36. Fountoulakis, M. & Langen, H. Identification of proteins by matrix-assisted laser desorption/ionization-mass spectrometry following in-gel digestion in low-salt, nonvolatile buffer and simplified peptide recovery. *Anal. Biochem.* **250**, 153–156 (1997).
37. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
38. Wootton, J. & Federhen, S. Statistics of local complexity in amino acid sequences and sequence databases. *Comput. Chem.* **17**, 149–163 (1993).
39. Labedan, B. & Riley, M. Gene products of *Escherichia coli*: sequence comparisons and common ancestries. *Mol. Biol. Evol.* **12**, 980–987 (1995).
40. Klein, P., Kanehisa, M. & DeLisi, C. The detection and classification of membrane-spanning proteins. *Biochim. Biophys. Acta* **815**, 468–476 (1985).
41. Frishman, D. & Argos, P. Seventy-five percent accuracy in protein secondary structure prediction. *Proteins* **27**, 329–335 (1997).
42. Nakashima, H., Nishikawa, K. & Ooi, T. The folding type of a protein is relevant to the amino acid composition. *J. Biochem.* **99**, 153–162 (1986).
43. Huynen, M. *et al.* Homology-based fold predictions for *Mycoplasma genitalium* proteins. *J. Mol. Biol.* **280**, 323–326 (1998).
44. Frishman, D. & Mewes, H. W. Pedantic genome analysis. *Trends Genet.* **13**, 415–416 (1997).
45. Bairoch, A. & Apweiler, R. The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1998. *Nucleic Acids Res.* **26**, 38–42 (1998).
46. Karp, P. D., Riley, M., Paley, S. M., Pellegrini-Toole, A. & Krummenacker, M. EcoCyc: Encyclopedia of *Escherichia coli* genes and metabolism. *Nucleic Acids Res.* **26**, 50–53 (1998).
47. Meyer, S. L. *Data Analysis for Scientists and Engineers*. (Wiley, New York, 1975).
48. Zhang, G. & Darst, S. A. Structure of the *Escherichia coli* RNA polymerase alpha subunit amino terminal domain. *Science* **281**, 262–266 (1998).
49. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296 (1991).

Acknowledgements

We thank K. Andersson for technical assistance and M. C. Bewley for help with Fig. 5b.

Correspondence and requests for materials should be addressed to F.U.H. (e-mail: uhartl@biochem.mpg.de).

Compaction as the origin of the unusual craters on the asteroid Mathilde

Kevin R. Housen*, Keith A. Holsapple† & Michael E. Voss*

* Shock Physics Group, MS 8H-05, The Boeing Company, PO Box 3999, Seattle, Washington 98124, USA

† Department of Aeronautics and Astronautics, 352400, University of Washington, Seattle, Washington 98195, USA

The asteroid Mathilde has suffered at least five giant impacts. Previous studies suggest that Mathilde's giant craters should be surrounded by blankets of ejecta that are kilometres deep^{1,2}, yet the craters show no evidence of filling by material excavated during later nearby impacts^{1,3}. Computer simulations of impacts have been used to suggest that the absence of ejecta arises because the impact energy is deposited in a small volume, due to Mathilde's unusually high porosity⁴, which produces ejecta velocities so high that nearly all of the material escapes Mathilde's gravitational field⁵. Here we report laboratory measurements of high-velocity impacts into porous material, which support an alternative explanation³: the crater is formed mainly by compaction, not excavation. The small amount of ejecta lofted in our experiments have velocities sufficiently low that nearly all of the material is redeposited within the crater bowl. The crater itself results from material being compressed, rather than ejected. This type of cratering implies that highly porous asteroids are minor contributors of meteorites, because essentially no material escapes the asteroids.

Numerical simulations of cratering face significant difficulties in realistically modelling the complex response of porous geological materials to high-speed impact. Therefore, it is useful also to study cratering experimentally. Laboratory experiments have a significant advantage in that they use actual geological materials, and provide benchmark data as tests for numerical simulations.

Experiments, however, necessarily involve craters much smaller than those of interest on Mathilde. To bridge this gap in size scale, a centrifuge can be used to simulate large-scale cratering events directly⁶. To show the basis for centrifuge modelling, consider the impact of a projectile of radius a , velocity U , and mass density δ , into an asteroid composed of a granular soil of strength s , mass density ρ , and friction angle ϕ , and whose surface gravitational acceleration is g' . Standard methods of dimensional analysis show that the dependence of crater diameter, D , on these seven parameters can be written in terms of five dimensionless groups^{6–8}:

$$\frac{D}{a} = f \left[\frac{g'a}{U^2}, \frac{s}{\rho U^2}, \frac{\rho}{\delta}, \phi \right]$$

Two impacts are physically the same (equivalent) if the four groups on the right side are the same for the two events. Then they must also have the same ratio of crater size to impactor size, D/a . If the two events involve the same materials and impact velocity, the sole remaining requirement for equivalence is that they have the same product $g'a$. The utility of increased gravity in centrifuge modelling stems from this condition: a big impact on an asteroid (large a , small g'), has the same value of $g'a$ as a small impact taking place in a centrifuge (small a , large g'). The centrifuge reproduces the physical conditions of the large event, but at greatly reduced size scale^{6–8}. In particular, it gives the same lithostatic stress as the large-scale event. Assuming that no significant variable has been omitted in the dimensional analysis, the small-scale test is physically the same as the large impact at the same velocity in the same material.

Equivalence also implies that $g'D$ is constant, because both $g'a$ and D/a are constant. Therefore, the diameter, D_C , of a centrifuge crater that simulates a crater of diameter D_M on a Mathilde-size body is given by $D_C = D_M g'_M / g'_C$, where g'_M and g'_C are the gravitational and centrifugal accelerations on Mathilde and the centrifuge, respectively. Mathilde's largest crater, Karoo, ($D_M = 33$ km, $g'_M = 1$ cm s⁻²) corresponds to a 6.7-cm diameter crater in a $g'_C = 500g$ experiment ($1g$ is 981 cm s⁻², so $g'_C = 4.9 \times 10^5$ cm s⁻²). Karoo and two other giant craters, Damodar and Kuznetsk, are shown in Fig. 1.

Centrifuge experiments are essential for complete simulations of impact events, even in cases where gravity has little effect on crater size, such as small impacts into a strong material like rock. This is because gravity controls the ballistics and final state of lofted material and, therefore, always affects the ejecta blanket. Fortunately, equivalence of crater size also guarantees that the ejecta blanket of a centrifuge crater is a geometric replica of the asteroid event².

We used a centrifuge to perform impact experiments at 500g into a low-density, porous, crushable silicate material (Fig. 2). Shot 1642 produced a remarkable crater that displayed essentially no ejecta outside the crater bowl. The small quantity that was ejected was less than 2% of the crater mass. This contrasts sharply with all previous experiments in soils, which always display well developed ejecta blankets².

The lack of ejecta was investigated further in shot 1648, an impact into the same material, but at 1g so that ejecta velocities could be measured from high-speed movies (initial ejection velocities are independent of gravity²). Although ejecta were produced in 1648, the speeds were quite low. The fastest observed ejecta had a speed of only 19 m s⁻¹, a greater portion were in the range 3–5 m s⁻¹, but most were below ~ 1 m s⁻¹. At 500g, an ejection speed of 5 m s⁻¹ gives a ballistic range of only 0.5 cm, or $\sim 20\%$ of the crater radius. Thus, nearly all ejecta in the 500g event must have landed inside the crater. The fact that the crater retained some 98% of its mass, yet was not filled by its own ejecta, can only be explained if most of the crater volume formed by compaction of pore spaces (Fig. 2), in contrast to the shearing and lofting observed in familiar geological materials. This same conclusion applies to the equivalent large craters on a Mathilde-size body, assuming its material behaves like our low-density, porous material.

Ejecta velocities were large enough to allow the 1g crater (1648) to develop a substantial ejecta blanket. Its 5-cm diameter is equivalent to a 50-m-diameter crater on a Mathilde-size body. This suggests that small craters on Mathilde could have ejecta blankets, although they probably would not be detected in the spacecraft images because of limited resolution.

Unlike the events reported here, cratering in common geological materials is largely an incompressible process. A crater is excavated as material is sheared and lofted to locations outside the crater. Thus, the stresses of the impact process must exceed both the shear strength (cohesion) of the material, and gravitational lithostatic stresses. This has led to two main categories of impact events, depending on whether crater size is determined mainly by material strength or by gravity. For impacts in which the cohesive strength exceeds the lithostatic stresses, the scaled crater diameter, D/a , is independent of event size. In the other case—that is, when lithostatic stresses are dominant ($\rho g'D \gg s$)—the scaled crater diameter diminishes with increasing event size^{7,8}.

The craters formed in our experiments fall into neither of these categories. The porous material was too fragile for strength measurements, but qualitative comparison with other weak materials⁹ indicate its strength is much smaller than the lithostatic stresses experienced in the 500g impacts ($\rho g'D = 3 \times 10^6$ dyn cm⁻²). That is, gravity played a more dominant role than strength. However, crater size was not controlled by gravity either; if it were, the 500g craters would have been substantially smaller than the 1g crater. In contrast, the 500g craters were in fact larger.

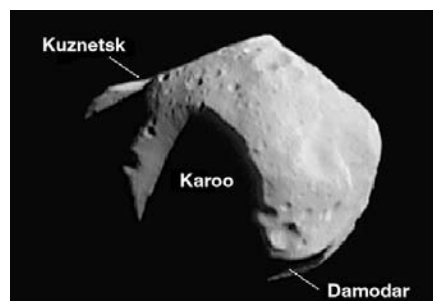


Figure 1 Mathilde, a 53-km-diameter asteroid imaged by the NEAR spacecraft. The largest crater, Karoo, has a diameter of 33.4 km. The two additional craters indicated

each have diameter of 29 km. A puzzling aspect of Mathilde is how these large craters could form in such close proximity and still retain a pristine appearance.

Rather than being controlled by shear strength or gravity, crater size in the porous material was determined mainly by the stress required to compact pore spaces. In loosely packed porous materials, compaction occurs at low pressures by breaking weak intergranular bonds and rearrangement of grains to reduce the volume of void spaces. This mechanism is distinct from the shock compaction of fully dense materials¹⁰, or the collapse of pores and cracks observed in silicate rocks¹¹, both of which require very high pressures, at least 10^9 dyn cm^{-2} . In contrast, tests in a hydraulic press showed that a pressure of $3 \times 10^7 \text{ dyn cm}^{-2}$ permanently compacted our porous material to twice its initial density, a density then comparable to that of dry fully dense sand.

The compaction stress was not the only factor that determined crater size; otherwise the volumes of the 1g and 500g craters would have been about equal. The movie at 1g showed a mass of slow

material lofted vertically to a height of $\sim 10 \text{ cm}$, which fell back into the crater. This material probably expanded, reducing the volume of the 1g crater. This bulking would not occur in large craters (or at 500g) because the ballistic height of the vertically launched material would be negligible compared to the crater size.

If our material is representative of bodies like Mathilde, these experiments show that the traditional strength- and gravity-dominated regimes of impact cratering do not pertain to porous asteroids. Instead, crater size is governed by pore compaction from the outgoing pressure shock. Consequently, the impact-driven evolution of porous asteroids may be entirely different from that of denser, rocky objects. On a rocky asteroid, impacts of all sizes eject debris, some of which escapes and may eventually strike the Earth as meteorites. The rest is re-deposited on the asteroid surface, degrading extant craters and contributing to

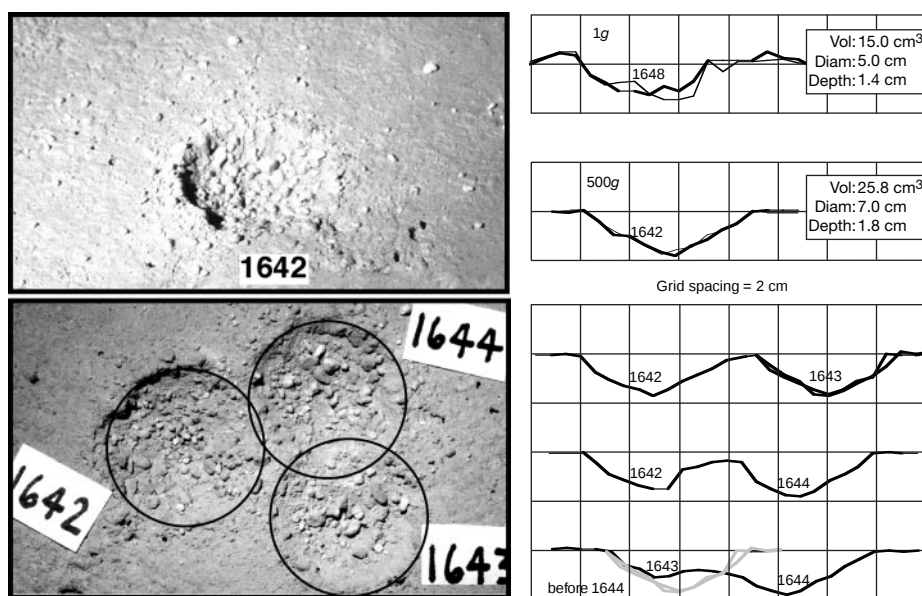


Figure 2 Experimental impact craters formed in a crushable material designed to illustrate the mechanisms that may occur on low-density asteroids. The target material has a bulk density of 0.9 g cm^{-3} , a porosity of 60%, and consists of a mixture of quartz sand, perlite (a porous, easily crushable, silicate), fly ash (a binding agent) and water. The perlite ranges from dust-sized particles up to the $\sim 5\text{-mm}$ chunks visible in the bottoms of the craters. Craters were formed by impacts of polyethylene cylinders (diameter, 0.65 cm; length, 0.63 cm; mass, 0.21 g; density, 1.04 g cm^{-3}) at 1.9 km s^{-1} . Shots 1642–1644 were performed at 500g on a centrifuge in order to simulate the lithostatic stress and ejecta ballistics of large cratering events on an asteroid. In crater 1642 (top left panel) only $\sim 2\%$ of the crater mass was ejected beyond the crater edge. This crater formed primarily by compaction, as opposed to excavation, of the target material. This was verified by imaging the regions under the crater with computed tomography, which showed an increase of density beneath the crater, with a maximum value nearly twice the initial

density. Compaction of pore spaces results in efficient damping of the shock as it propagates into the target, which may explain why the large craters on Mathilde formed in such close proximity with little evidence of seismic disturbance of nearby craters. To investigate this further, two additional craters at 500g were formed close to the 1642 crater (bottom left panel). As shown in the crater profiles (right panels), formation of crater 1643 had no visible effect on crater 1642, even though their rims were nearly touching. Crater 1644 formed even closer, breaching the rim of craters 1642 and 1643. Intersection of the rims resulted in some slumping of material into craters 1642 and 1643. These experiments show that large craters in porous crushable targets can form with little degradation of existing proximal craters. The appearance of Mathilde's pristine large craters is probably due to both suppression of ejecta and damping of the shock; both of which are a consequence of a highly porous material.

regolith build-up. In contrast, on a highly porous asteroid, only small impacts (perhaps $D < 1$ km for Mathilde) produce ejecta deposits outside the crater rim (like shot 1648), whereas blankets would be absent around large craters. Only a small amount of ejecta would escape Mathilde. Such asteroids would liberate significant meteoritic material only from catastrophic impacts that shatter and disperse the whole body.

High porosity does not guarantee formation of compaction craters. For example, dry sand has a porosity of 35%, but sand craters form primarily by excavation, with significant ejecta blankets at all sizes. Compaction in sand is minimal because it is already near a 'fully dense' state (the most efficient packing of particles). In this case, compaction cratering could only occur by crushing of the constituent sand grains, which requires stresses much higher than those experienced by most of the cratered material. Most granular silicate materials are at their fully dense state when their bulk density is in the range $2\text{--}3\text{ g cm}^{-3}$. Thus, compaction cratering in silicates can only occur if the bulk density is well below $\sim 2\text{ g cm}^{-3}$. We note that large craters on the martian moons Phobos and Deimos (densities $\sim 1.9\text{ g cm}^{-3}$) do not show strong evidence of compaction effects¹², probably because they are close to the fully dense state. Furthermore, even initially highly porous asteroids about ten times larger than Mathilde's diameter would have lithostatic stresses comparable to the crush pressure of the material used here, and would naturally compact to near a fully dense state due to self-gravity. Therefore, compaction cratering is not expected to be common on large asteroids.

High porosity may even be a fleeting characteristic of Mathilde-sized asteroids. As shown by laboratory experiments^{13,14}, and by Mathilde, highly porous bodies can withstand multiple large impacts without disruption. Each impact locally compresses the asteroid, because its volume decreases by the crater volume, while all mass is retained. Formation of the five largest craters on Mathilde ($D_M > 20$ km), increased its bulk density by $\sim 20\%$. Hence, Mathilde's initial density may have been even lower than the present value, especially considering that additional large craters may exist on the unobserved half of its surface. Over time, porous bodies may be compacted by impacts to the point of being fully dense. Ejection velocities would then increase, allowing escape of some debris and formation of ejecta blankets around large craters, much as we expect for compact, rocky bodies. □

Received 12 April; accepted 17 September 1999.

1. Veverka, J. *et al.* NEAR's flyby of 253 Mathilde: Images of a C asteroid. *Science* **278**, 2109–2112 (1997).
2. Housen, K. R., Schmidt, R. M. & Holsapple, K. A. Crater ejecta scaling laws: Fundamental forms based on dimensional analysis. *J. Geophys. Res.* **88**, 2485–2499 (1983).
3. Chapman, C. R., Merline, W. J. & Thomas, P. Cratering on Mathilde. *Icarus* (submitted).
4. Veverka, J. *et al.* NEAR Encounter with asteroid 253 Mathilde: Overview. *Icarus* (submitted).
5. Asphaug, E. *et al.* Disruption of kilometre-sized asteroids by energetic collisions. *Nature* **393**, 437–440 (1998).
6. Schmidt, R. M. & Holsapple, K. A. Theory and experiments on centrifuge cratering. *J. Geophys. Res.* **85**, 235–252 (1980).
7. Holsapple, K. A. & Schmidt, R. M. Point source solutions and coupling parameters in cratering mechanics. *J. Geophys. Res.* **92**, 6350–6376 (1987).
8. Holsapple, K. A. The scaling of impact processes in planetary sciences. *Annu. Rev. Earth Planet. Sci.* **21**, 333–373 (1993).
9. Housen, K. R., Schmidt, R. M. & Holsapple, K. A. Laboratory simulations of large scale fragmentation events. *Icarus* **94**, 180–190 (1991).
10. Kieffer, S. W. From regolith to rock by shock. *Moon* **13**, 301–320 (1975).
11. Feves, M., Simmons, G. & Siegfried, R. W. in *The Earth's Crust. Its Nature and Physical Properties* (ed. Heacock, J. G.) (Geophysical Monogr. 20, American Geophysical Union, Washington, DC, 1977).
12. Thomas, P. C. Ejecta emplacement on the martian satellites. *Icarus* **131**, 78–106 (1998).
13. Love, S. G., Hörz, F. & Brownlee, D. E. Target porosity effects in impact cratering and collisional disruption. *Icarus* **105**, 216–224 (1993).
14. Stewart, S. T. & Ahrens, T. J. Porosity effects on impact processes in solar system materials. Vol. 30, CD-ROM (*Lunar Planetary Science Conference*, Houston, Texas, 1999).

Acknowledgements

This work was supported by NASA's Planetary Geology and Geophysics Program.

Correspondence and requests for materials should be addressed to K.R.H. (e-mail: kevin.r.housen@boeing.com).

Spatial sampling of crystal electrons by in-flight annihilation of fast positrons

A. W. Hunt^{*†}, D. B. Cassidy^{*†}, F. A. Selim[‡], R. Haakenaasen[§], T. E. Cowan[†], R. H. Howell[†], K. G. Lynn^{||} & J. A. Golovchenko^{*¶}

^{*} Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA

[†] Lawrence Livermore National Laboratory, Livermore, California 94550, USA

[‡] Engineering Physics Department, Alexandria University, Alexandria 21544, Egypt

[§] Norwegian Defense Research Establishment, 2027 Kjeller, Norway

^{||} Department of Physics, Washington State University, Pullman, Washington 99164, USA

[¶] Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

[#] The Rowland Institute for Science, Cambridge, Massachusetts 02142, USA

Energetic, positively charged particles travelling along a low-index crystal direction undergo many highly correlated, small-angle scattering events; the effect of these interactions is to guide or 'channel' (refs 1–8) the particles through the lattice. Channeling effectively focuses positive particles into the interstitial regions of the crystal: nuclear collisional processes such as Rutherford backscattering are suppressed, while the number of interactions with valence electrons increases. The interaction of channelled positrons with electrons produces annihilation radiation that can in principle^{9–12} serve as a quantitative, spatially selective probe of electronic charge and spin densities within the crystal: in the interstitial regions, two-photon annihilation is enhanced relative to single-photon annihilation, because the latter process requires a nuclear recoil to conserve momentum. Here we report observations of single- and two-photon annihilation from channelled positrons, using a monoenergetic beam flux of 10^5 particles per second. Comparison of these two annihilation modes demonstrates the ability of channelled positrons to selectively sample valence electrons in a crystal. Useful practical implementation of the technique will require the development of more intense positron beams with fluxes approaching 10^7 particles per second.

These experiments were performed at the recently constructed 3-MeV monoenergetic positron beamline at Lawrence Livermore National Laboratory. An electrostatic accelerator was modified to accommodate a 109 mCi ^{22}Na positron source, a tungsten moderator, as well as appropriate focusing optics in the terminal. The positrons were accelerated to 2.65 MeV and then passed through two bending magnets, two solenoid focusing lenses and two quadrupole astigmatism correctors, all designed, constructed or modified for

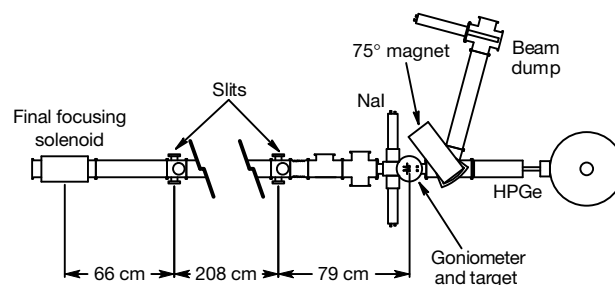


Figure 1 Experimental set-up. Shown is a diagram of the final leg of the 3-MeV monoenergetic positron beamline at Lawrence Livermore National Laboratory.

this experiment. The beam was then delivered to the stretch of optics shown in Fig. 1, where the channelling experiments were performed.

The positrons passed through a final solenoid focusing lens and a pair of collimation slits that defined the angular divergence and spatial size of the beam as it impinged on the target. Typical beam characteristics on target were 200,000 positrons per second in a spot 3 mm in diameter, with a maximum angular divergence of 0.25° . The energy spread of the beam was less than 0.4% of the full beam energy. The target was a 0.6- μm -thick free-standing gold crystal oriented in the (110) direction grown at the University of Aarhus with a surface minimum yield of 3% for 2 MeV α -particles. It was housed in a two-axis goniometer with a scintillator mounted on the front side. The scintillator had a hole 4.5 mm in diameter aligned with the target and was used for beam optimization and background subtraction.

A large-angle scattering scintillator was placed behind the target that detected positrons scattered between 10° and 20° . Further downstream a 75° bending magnet deflected the transmitted positrons to a well shielded beam dump. This was a specially designed double-focusing magnet with a 10° vertical and horizontal acceptance angle. Multiple scattering measurements and calculations predicted an angular divergence of half this value, ensuring efficient collection of the transmitted beam¹³. A scintillation detector used for normalization was positioned in the focal plane of the 75° magnet.

A multiple γ -ray spectroscopy system, consisting of a high purity germanium (HPGe) detector 7.3 cm in diameter and two 3-inch NaI scintillators, recorded annihilation photons from the gold crystal. The HPGe detector was placed in the forward direction 37.5 cm from the target and subtended a half angle of 5.6° . Annihilation photons with energies in excess of 3.3 MeV were constrained to emit in this forward cone by momentum conservation. The two NaI scintillators were placed on either side of the target chamber 8.9 cm in front of the target, collecting photons emitted between 142° and 163° . These detectors recorded the ~ 300 keV photons emitted in coincidence with the forward ~ 3.37 MeV γ -rays from two-photon annihilation. The signals from all detectors were routed to a multiple-parameter data acquisition system that recorded energy and time information for each detector on an event by event basis. Low background two-photon annihilation energy spectra were obtained by requiring a coincidence within 30 ns between the HPGe detector and one NaI scintillator. To identify single-photon annihilation events, a spectrum from the HPGe detector was collected concurrently with no coincidence requirement.

Figure 2 shows energy spectra obtained from the HPGe detector without (curve A) and with (curve B) the coincidence requirement. These particular demonstration spectra were obtained using a 3.6- μm -thick amorphous gold foil for enhanced statistics. In the single-photon annihilation process a monoenergetic γ -ray is emitted with the nucleus recoiling to conserve momentum. The energy of the photon, E_s , is given by

$$E_s = 2mc^2 + T_+ - |E_b|$$

where mc^2 is the rest mass energy of the electron, T_+ is the positron kinetic energy and E_b is the electron binding energy. Single photon annihilation is easily identified in curve A (that is, no coincidence required) of Fig. 2. The K-shell ($|E_b|=81$ keV), L-shell ($|E_b|=13$ keV) and M-shell ($|E_b|=2.6$ keV) annihilation events are seen at 3,591 keV, 3,659 keV and 3,670 keV respectively. The requirement of a nuclear recoil severely restricts this annihilation process to the most tightly bound electrons¹⁴.

Two-photon annihilation does not require a nuclear recoil and can therefore occur with all electrons in the target material. The two emitted photons have energies that depend strongly on the emission angles. For 2.65-MeV positrons the photon energies range from 276

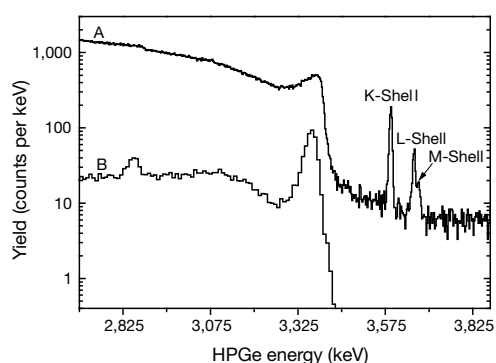


Figure 2 HPGe energy spectra. Annihilation radiation results from 2.65-MeV positrons impinging on a 3.6- μm polycrystalline gold foil. Curve A, no coincident photon required. Curve B, coincident photon required within 30 ns in one of the 3-inch detectors.

keV for 180° emission to 3,396 keV for 0° emission. The edge seen in Fig. 2 at 3,390 keV is from the most forward directed of these photons. By requiring a coincidence with the NaI scintillator the background is dramatically reduced and two-photon annihilation is clearly distinguished, as seen in curve B of Fig. 2. The peak at 3,370 keV is the signature of the forward-detected γ -rays from two-photon annihilation. The full-width at half-maximum of 46.6 keV is 6 times larger than the energy resolution of the HPGe detector and is determined by the kinematics of the two-photon annihilation process in conjunction with the angular acceptances of the detectors. The Compton edge at 3,130 keV and the continuous plateau at lower energy are due to γ -ray Compton scattering out of the HPGe detector. The small peak at 2,859 keV is the single γ -ray escape peak from recombining positron-electron pairs created in the detector. Both of these mechanisms allow valid two-photon annihilation events to be recorded below the peak at 3,370 keV. Including the events from ~ 1 MeV through the photo peak significantly improves the statistics in our channelling studies.

Using the scattering detector and spectra similar to those in Fig. 2, the angular yields from large-angle scattering, two-photon annihilation and K-shell one-photon annihilation were measured for positrons channelled through the 0.6- μm -thick gold crystal near a (110) axial direction. (With this crystal thickness, the positron flux was too weak to produce observable L-shell or M-shell one-photon annihilation signals.) The results are plotted in Fig. 3. The horizontal axis represents the angle between the positron beam and the (110) axis of the crystal. The vertical axis is the observed yield normalized to unity for a non-channelling direction. Each annihilation datum took at least 12 hours to accumulate. For scattered positrons the on-axis minimum yield is 6.5% and the full-width at half-minimum of the channelling dip is 1.65° . The on-axis reduction of large-angle nuclear scattering reflects the strong channelling suppression of the small-impact parameter collisions ($b < 4 \times 10^{-3}$ Å) required for deflection to occur into the scattering detector. The observed width is 8% below the calculated value of $\sim 1.8^\circ$ obtained from a dynamical diffraction yield calculation³ which does not account for dechannelling processes. This calculation is based on a formalism that successfully describes the well known phenomenon of electron-channelling radiation¹⁵. Inclusion of dechannelling processes into the computation would improve the agreement but is a complex task.

The similarity between the channelling yield curves for K-shell annihilation and large-angle scattering may be understood by considering the K-shell Bohr radius ($a = 7 \times 10^{-3}$ Å), the scattering impact parameter ($b = 4 \times 10^{-3}$ Å) and the gold atomic thermal vibration amplitude ($\rho = 0.12$ Å). Since both processes require close proximity to the nucleus, smaller than the thermal vibration

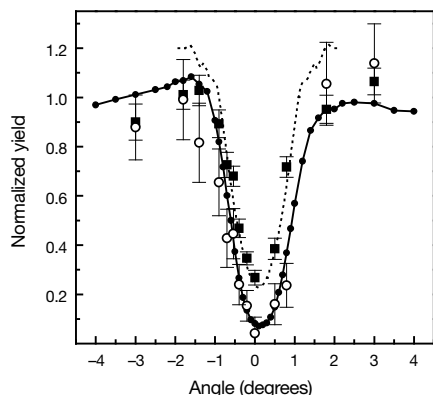


Figure 3 Normalized angular yield curves. Results are shown for 2.65-MeV positrons undergoing large-angle scattering (solid line with filled circle), single-photon annihilation (open circle) and two-photon annihilation (filled square). Angle on horizontal axis measured from the (110) axis of the 0.6- μm gold crystal. The dashed line is a dynamical diffraction yield calculation for two-photon annihilation including a dechannelling correction from the experimentally observed scattering minimum yield.

amplitude, they are expected to exhibit very similar channelling characteristics. In comparison, the sensitivity of two-photon annihilation to the outer and valence electrons causes a striking difference. The full width is reduced to $1.25 \pm 0.09^\circ$ and the minimum yield is four times higher at $27 \pm 3\%$. This is unambiguously attributed to the extended nature of the electron clouds around the nuclei and the delocalized electrons in the interstitial region. Since the single-photon annihilation cross-section for these electrons is negligible, the smaller minimum yield of single- versus two-photon annihilation verifies that annihilation with the outer and valence electrons is enhanced by the channelling process.

Also shown in Fig. 3 is the dynamical diffraction yield calculation for two-photon annihilation. The electron densities and crystalline potential were calculated using a state-of-the-art augmented plane wave method¹⁶. The calculated full width agrees well with the two-photon annihilation data and the calculated minimum yield is 14.5% in the absence of dechannelling. Dechannelled positrons are not confined to the interstitial regions and consequently they may interact with the core electrons which increases the annihilation yield. From the observed scattering minimum yield we can estimate that dechannelling will increase the two-photon minimum yield to 23%, which is 1.5 standard deviations from the experimentally observed value. Using these calculations we can also estimate the valence-electron contribution. In gold, we consider the valence electrons to be the 10 electrons in the 5d shell and the electron in the 6s shell. At large incident angles, where channelling does not occur, positrons penetrate into all regions of the crystal and therefore sample all electrons equally. Since there are 11 valence electrons, they constitute 14% of the total sampled at large angles. In contrast, channelled positrons are focused into the interstitial region, where the valence electrons account for almost all of the electron density and their annihilation contribution increases to 56% at minimum yield with the experimentally observed dechannelling correction and background subtraction included. An implication of this enhancement is the potential to investigate charge density effects reflecting crystal bonding. For gold, our calculations predict a 3.5% increase in minimum yield for the case of a bonded crystal compared with a pseudo-crystal comprised of neutral unbonded atoms. This effect would be easily measurable with an increase in positron flux of 2 or 3 orders of magnitude.

Thus the enhanced minimum yield observed for two-photon over one-photon annihilation confirms our expectations that two-photon annihilation channelling effects would be very sensitive to

electron densities in the interstitial regions of crystals. Indeed, the relative influence of valence electrons on channelling annihilation can be estimated to be about ten times larger than the relative contribution of valence electrons to low-order Fourier components of the charge density. These Fourier components are directly measured by X-ray diffraction techniques and have been used to determine electron densities in crystals^{17,18}. These components have also been obtained by dynamical electron diffraction methods^{19,20}, which have an advantage over X-ray methods in that the core-electron contribution to the scattering is cancelled by the nuclear potential. The spatial probe provided by positron-annihilation channelling has an enhanced sensitivity to valence electrons that is determined by the positive charge of the probe. It would be interesting to determine whether a similar advantage exists for positrons in dynamical diffraction transmission studies.

If the positron-channelling phenomena revealed by this research are eventually to result in an accurate, easily applied probe of electrons in solid-state materials, it is clear that thinner crystals are required in order to reduce dechannelling effects and high positron-beam currents are needed in order to perform experiments on a practical timescale; monoenergetic MeV positron-beam currents greater than 10^7 positrons per second will be required. Stronger positron sources, improved moderators, polarization control, and more efficient detection and beam transport capabilities are possible with current technology. This new capability will also enable significant extensions of the experimental approach reported here. A next-generation beam facility will allow the development of practical atomic-scale channelling measurements of electronic spin densities^{9–12}, and momentum profiles in addition to valence and bonding electron density maps. We particularly look forward to the prospect of performing spin-density measurements on thin-film magnetic crystals $\sim 1,000 \text{ \AA}$ in thickness. Similar measurements on such small samples are well beyond the capabilities of neutron diffraction techniques. □

Received 29 June; accepted 20 September 1999.

1. Lindhard, J. Motion of swift charged particles, as influenced by strings of atoms in crystals. *Phys. Lett.* **12**, 126–128 (1964).
2. Lindhard, J. Influence of crystal lattice on motion of energetic charged particles. *K. Danske Vidensk. Selsk. Mat.-fys. Medd.* **34**(14), 1–64 (1965).
3. Howie, A. Diffraction channeling of fast electrons and positrons in crystals. *Phil. Mag.* **4**, 223–237 (1966).
4. Uggerhøj, E. Orientation dependence of the emission of positrons and electrons from ^{64}Cu embedded in single crystals. *Phys. Lett.* **22**, 382–383 (1966).
5. Uggerhøj, E. & Andersen, J. U. Influence of lattice structure on motion of positrons and electrons through single crystals. *Can. J. Phys.* **46**, 543–550 (1968).
6. Andersen, J. U., Augustyniak, W. M. & Uggerhøj, E. Channeling of positrons. *Phys. Rev. B* **3**, 705–711 (1971).
7. Andersen, J. U., Andersen, S. K. & Augustyniak, W. M. Channeling of electrons and positrons. *K. Danske Vidensk. Selsk. Mat.-fys. Medd.* **39**(10), 1–58 (1977).
8. Schultz, P. J., Logan, L. R., Jackman, T. E. & Davies, J. A. Channeling effects on positron transmission through thin crystals. *Phys. Rev. B* **38**, 6369–6379 (1988).
9. Haakenaasen, R. *et al.* Quantum channeling effects for 1 MeV positrons. *Phys. Rev. Lett.* **75**, 1650–1653 (1995).
10. Hau, L. V. *et al.* On the role of classical and quantum notions in channeling and the development of fast positrons as a solid state probe of valence electron and spin densities. *Nucl. Instrum. Methods B* **119**, 30–41 (1996).
11. Hau, L. V., Hunt, A. W., Golovchenko, J. A., Haakenaasen, R. & Lynn, K. G. in *Proc. 11th Int. Conf. Positron Annihilation* (eds Jean, Y. C., Eldrup, M., Schrader, D. M. & West, R. N.) 119–123 (Trans Tech, Uetikon, Zurich, Switzerland, 1997).
12. Haakenaasen, R. *et al.* in *Proc. 11th Int. Conf. Positron Annihilation* (eds Jean, Y. C., Eldrup, M., Schrader, D. M. & West, R. N.) 257–259 (Trans Tech, Uetikon, Zurich, Switzerland, 1997).
13. Selim, F. A. *Penetration of Energetic Positrons through Crystalline and Noncrystalline Media*. Thesis, Alexandria Univ. (1999).
14. Palathingal, J. C., Asoka-Kumar, P., Lynn, K. G. & Wu, X. Y. Nuclear-charge and positron-energy dependence of the single-quantum annihilation of positrons. *Phys. Rev. A* **51**, 2122–2130 (1995).
15. Andersen, J. U., Bonderup, E., Lægsgaard, E., Marsh, B. B. & Sørensen, A. H. Axial channeling radiation from MeV electrons. *Nucl. Instrum. Methods* **194**, 209–224 (1982).
16. Rajagopal, A. K. in *Advances in Chemical Physics* (eds Prigogine, I. & Rice, S. A.) Vol. 41 (John and Sons, New York, 1980).
17. Dawson, B. The covalent bond in silicon. *Proc. R. Soc. Lond. A* **298**, 379–394 (1967).
18. Aldred, P. J. E. & Hart, M. The electron distribution in silicon. *Proc. R. Soc. Lond. A* **332**, 223–238 (1973).
19. Zuo, J. M., Kim, M., O'Keeffe, M. & Spence, J. Direct observation of d-orbital holes and Cu-Cu bonding in Cu_2O . *Nature* **401**, 49–52 (1999).

20. Zuo, J. M., O'Keeffe, M., Rez, P. & Spence, J. C. H. Charge density of MgO: Implications of precise new measurements for theory. *Phys. Rev. Lett.* **78**, 4777–4780 (1997).

Acknowledgements

We thank J. Chevalier, W. Gibson, R. A. Levesque, J. U. Andersen, M. Weinert and J. C. Palathingal for contributions to this research. This work was supported by the National Science Foundation, the Norwegian Research Council and the Department of Energy through Lawrence Livermore National Laboratory.

Correspondence and requests for materials should be addressed to J.A.G.

Diffusion mechanisms in metallic supercooled liquids and glasses

X.-P. Tang*, Ulrich Geyer†, Ralf Busch‡, William L. Johnson‡ & Yue Wu*

* Department of Physics and Astronomy, University of North Carolina, Chapel Hill, North Carolina 27599-3255, USA

† Erstes Physikalisches Institut der Universität Göttingen, Bunsenstraße 9, D-37073 Göttingen, Germany

‡ W.M. Keck Laboratory of Engineering Materials, California Institute of Technology, Pasadena, California 91125, USA

The mechanisms of atomic transport in supercooled liquids and the nature of the glass transition are long-standing problems^{1–4}. Collective atomic motion is thought to play an important role^{4–6} in both phenomena. A metallic supercooled liquid represents an ideal system for studying intrinsic collective motions because of its structural similarity to the “dense random packing of spheres” model⁷, which is conceptually simple. Unlike polymeric and network glasses, metallic supercooled liquids have only recently become experimentally accessible, following the discovery of bulk metallic glasses^{8–12}. Here we report a ⁹Be nuclear magnetic resonance study of Zr-based bulk metallic glasses^{8,9} in which we investigate microscopic transport in supercooled liquids around the glass transition regime. Combining our results with diffusion measurements, we demonstrate that two distinct processes contribute to long-range transport in the supercooled liquid state: single-atom hopping and collective motion, the latter being the dominant process. The effect of the glass transition is clearly visible in the observed diffusion behaviour of the Be atoms.

In the past few years, bulk metallic glasses with extraordinary glass-forming abilities have been discovered, such as Zr_{41.2}Ti_{13.8}Cu_{12.5}Ni₁₀Be_{22.5} (Vit1; ref. 8) and Zr_{46.75}Ti_{8.25}Cu_{7.5}Ni₁₀Be_{27.5} (Vit4; ref. 9). Their crystallization kinetics in the supercooled liquid state (SLS) is very slow compared to conventional metallic glasses¹⁰, so that investigations of atomic transport in the metallic SLS become possible. For Vit1 and Vit4, the long-range atomic transport has been investigated by diffusion measurements^{13–17}. It is, however, difficult to study atomic motion at microscopic scales in metallic systems. The conventional probes of local structural relaxation, such as dielectric measurement, are not applicable. For nuclear magnetic resonance (NMR), the difficulty arises from the fact that the relevant timescale of atomic motion near the glass transition is very long compared to the typical nuclear spin–lattice relaxation time (T_1) in metals. Fortunately, this is not the case for ⁹Be NMR in Vit1 and Vit4 (T_1 is about 4 s at 300 K for ⁹Be). It was shown recently^{18,19} that the ⁹Be spin alignment echo technique (SAE) is very effective in detecting Be hopping in glassy Vit1 and Vit4 with atomic jump rates as low as 0.1 s^{−1}.

Figure 1 shows the temperature dependence of the interdiffusion coefficients D of Be in Vit1 and Vit4. Below 620 K, which lies in the calorimetric glass-transition regime, Be diffusion shows Arrhenius behaviour with activation energies of about 1.1 eV in both alloys^{13,14}.

At higher temperatures, the temperature dependence of Be diffusion increases significantly, resulting in kinks in the Arrhenius plot of D . A kink has also been observed in $D(T)$ data for other elements, including Ni (ref. 15), B and Fe (ref. 16), and Co (refs 16, 17) in Vit4. One explanation of the observed kink suggests¹⁵ that the diffusivity with larger activation energy above the kink temperature represents the intrinsic property of the SLS. The smaller activation energy below the kink temperature is attributed to an incomplete structural relaxation of the glassy state that would disappear after sufficiently long annealing. An alternative explanation¹³ assumes that the observed low-temperature diffusion behaviour is an intrinsic property of the relaxed glassy state. Diffusion both above and below the kink temperature is attributed to hopping of single atoms of Be. This model explains the higher apparent activation energy above 620 K in terms of structural changes of the matrix in the SLS on the timescale comparable to the time interval between Be hopping events. It was assumed that such structural changes would enhance the probability of Be hopping and lead to increased diffusion¹³. Clearly, to elucidate the nature of atomic motion, information at the microscopic scale is needed.

The SAE technique^{18–20} is based on the Jeener–Broekaert sequence $90^\circ_x - \tau_1 - 45^\circ_y - \tau_2 - 45^\circ_z - \tau_3$, where 90° and 45° are the tipping angles of the radio-frequency pulses, x and y are the phases of the pulses, and τ_1 , τ_2 and τ_3 are delays between pulses. Here the first two pulses with a short τ_1 (6–15 μ s) create a pure quadrupole order, and a spin alignment echo is formed at $\tau_2 = \tau_1$ following the third pulse. Under the condition $\tau > T_2$ (the ⁹Be spin–spin relaxation time T_2 is about 1.5 ms), the echo amplitude is proportional to the ensemble-averaged correlation function $f(\tau) = \langle \sin(\omega\tau_1) \sin(\omega'\tau_1) \rangle \exp(-\alpha\tau/T_1)$. Here α is a constant and is equal to 3 as measured in Vit1 and Vit4^{18,19}. The frequencies ω and ω' , varying over a range of $2\pi \times 100$ kHz in Vit1 and Vit4¹⁹, are determined by the truncated quadrupole interactions experienced by the spin during τ_1 and τ_2 , respectively^{18–20}. The truncated quadrupole interaction with respect to the external magnetic field differs from site to site because of the disordered atomic structure of glasses and materials in a SLS. Thus, hopping of a Be atom is expected to change ω randomly within its range, and can be detected by the SAE technique through $f(\tau)$ decay. On the other hand, the SAE technique is not sensitive to collective motion involving Be atoms and most of their nearest neighbours. The truncated quadrupole interaction at the Be site changes very little under collective translation. This technique is also insensitive to constrained collective rotation with small rotation angle because $(\omega - \omega')\tau_1 \ll 1$. In both Vit1 and Vit4, a single-exponential

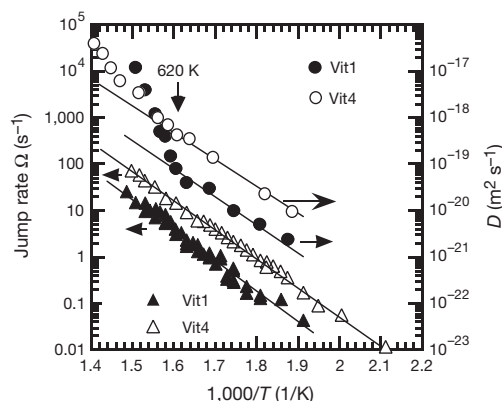


Figure 1 Elastic back-scattering of 6.5 MeV ⁴He²⁺ ions and SAE measurements in bulk metallic glasses. The temperature dependence of Be diffusion coefficients D (measured by elastic back-scattering) and atomic jump rates Ω (measured by SAE) are shown for Vit1 and Vit4. D (scale on the right) and Ω (scale on the left) are plotted versus $1,000/T$. A kink at 620 K, which is around the calorimetric glass-transition regime, is observed in the $D(T)$ curves in both Vit1 and Vit4. This kink is absent in the $\Omega(T)$ curves, which follow the same Arrhenius behaviour in both the SLS and glassy state.

decay over nearly three orders of magnitude is observed¹⁹ for the alignment echo as a function of τ . Such behaviour is expected for Be motions causing random changes of ω . In such a case, $f(\tau)$ is then proportional to $\exp(-\tau/T_{QE})$, where $1/T_{QE}=3/T_1+\Omega$, and Ω is the atomic jump rate^{18,19}. The single-exponential decay measured by the SAE technique is inconsistent with constrained jumps between neighbouring sites such as double-well potentials. In such cases, $f(\tau)$ should decay non-exponentially and should have a residual intensity^{20,21} which decays with the rate $3/T_1$ rather than $1/T_{QE}$. The residual intensity results from the finite probability of an atom returning to the original configuration after many jumps. The measured exponential decay as found by SAE is consistent with unconstrained Be hopping assisted by the thermal fluctuation of spread-out free volume¹⁹. In glasses the free volume is distributed among all atoms. Redistribution of free volume through thermal fluctuation leads to temporary formation of volumetric defects and induces atomic diffusion.

Spin-lattice relaxation of the quadrupole order induces alignment echo decay through transitions among spin states, and is sensitive to fast motion of the order of the ⁹Be Larmor frequency (56.2 MHz). This contribution can be investigated independently by measurement of T_1 . In Vit1 and Vit4, $1/T_1$ is proportional to T over the entire temperature range measured^{18,19}. This reveals the origin of $1/T_1$ as Korringa relaxation, due to the hyperfine interaction between the conduction electron spins and nuclear spins, consistent with the measured value of $\alpha=3$. Thus, the effect of fast motion on the timescale of the Larmor frequency is negligible for SAE. Furthermore, T_2 and ⁹Be spectra are also temperature-independent, indicating that motions on timescales of 1 ms and shorter are also insignificant.

The measured Ω values follow Arrhenius behaviour in both Vit1 and Vit4 (Fig. 1). Fitting the measured Ω using $\Omega(T)=\Omega_0\exp(-E_a/k_B T)$ yields an activation energy $E_a=(1.2\pm0.15)$ eV for both Vit1 and Vit4. Within the experimental uncertainty, this E_a is identical to those obtained by diffusion measurements below the kink temperature. Also, Ω in Vit4 is larger than that in Vit1, consistent with the trend of the diffusion constant below the kink temperature. These observations suggest again that the origin of the motion observed by SAE is the same as that measured by diffusion below the kink temperature. If the origins of the motions detected by SAE and diffusion are different, the measured diffusion below 620 K has to be due to collective motions to escape SAE detection.

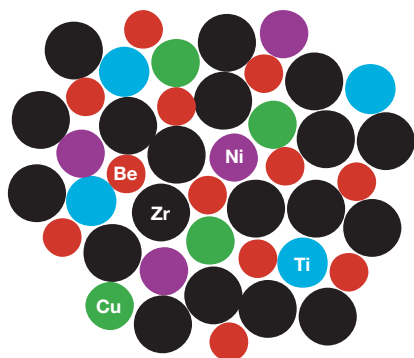


Figure 2 A two-dimensional illustration of randomly and densely packed spheres. The radius of the sphere labelled by a certain element is proportional to the radius of the corresponding atom. Single-atom hopping of atoms is assisted by thermal fluctuation of spread-out free volume^{4,19}. Thermal fluctuation of spread-out free volume leads to the temporary formation of a free volume larger than a critical volume near the Be atom. Hopping of Be into such temporarily formed free volume leads to Be diffusion. After hopping, Be resides at the other side of its previous nearest neighbours. Using 2.2, 2.5, 2.9 and 3.2 Å as the diameters of Be, Ni (Cu), Ti and Zr, respectively, the typical Be hopping distance of each hopping step is estimated to be around 3–5 Å. For a three-dimensional structure of dense random packing, the estimated typical Be hopping distance of each hopping step is approximately the same.

For collective motion, similar diffusion behaviour should be expected for Be and other larger atoms. However, Be atoms are observed^{13–17} to diffuse significantly faster than other larger atoms below 620 K. Therefore, the observed Be diffusion below the kink temperature is via single-atom hopping which is detectable by both diffusion and SAE. For this mechanism, based on the equation $D=f\Omega\lambda^2/6$ where f is the correlation factor ($0 < f \leq 1$) and λ^2 is the mean square displacement per jump, $f\lambda^2$ values can be derived from data shown in Fig. 1. They are $(4.5 \text{ Å})^2$ and $(2.8 \text{ Å})^2$ for Vit1 and Vit4, respectively, and are consistent with the estimated hopping distance of 3–5 Å (Fig. 2).

In the present work, measurements using the SAE technique have been extended above the kink temperature. Surprisingly, no kink is observed in the Arrhenius plot of Ω (Fig. 1). The measured Ω in the SLS obeys the same Arrhenius behaviour as that below 620 K in both Vit1 and Vit4. This demonstrates that single-atom Be hopping is an intrinsic property of both the SLS and the relaxed glassy state. Since this process is present at temperatures well into the SLS, its relatively low activation energy cannot be a result of incomplete structural relaxation of the glassy state. In addition, SAE can determine the hopping rate at a given temperature in a few minutes (ref. 19). This allows the measurement of Ω as a function of isothermal annealing time. No annealing-time dependence of Ω is observed even below 620 K. Despite the agreement between the results of the diffusion and SAE measurements below the kink temperature, the rapid increase of Be diffusion above 620 K is missing in the present SAE result. This proves that the enhanced diffusion above the kink temperature is not due to the increase of the Be hopping rate as suggested previously¹³. Such a rate increase should also cause an enhancement of Ω , and this is not the case.

The combined SAE and diffusion studies indicate that the long-range transport of Be atoms in the SLS is due to two diffusion processes, one of which is single-atom hopping. The other process, which dominates in the SLS, does not significantly change the truncated quadrupole interactions at most Be sites over the average time interval between two hopping events. This suggests that this process is collective, involving most of the nearest neighbours of Be atoms. The size of the collectively moving group of atoms cannot be evaluated at this point without theoretical calculations of quadrupole interactions. The estimated time interval between two consecutive Be jumps is around 330 ms in Vit1 and 80 ms in Vit4 at 620 K, according to SAE measurement. These are about a factor of 100 shorter than the macroscopic structural relaxation time as determined by shear viscosity¹⁰. Thus, only a small fraction of elementary diffusion events contribute to macroscopic structural relaxation, consistent with the breakdown of the Stokes–Einstein relation between macroscopic flow and diffusion in Vit1 and Vit4^{13,14}. This suggests that the number of collectively moving atoms is small. There are other indications which support the idea of atomic clusters. For instance, the small isotope effect of Co diffusion in several metallic glasses and in the SLS has been attributed to highly cooperative hopping of a group of about ten atoms^{4,17}. We note that although the Be atom is relatively small compared to atoms of other elements in Vit1 and Vit4, the effect of the glass transition is clearly seen in the measurements of Be diffusion. In contrast, diffusion of small molecules in some polymeric glasses is decoupled from the glass transition²². This is in line with the dense packing of atoms in Vit1 and Vit4, and demonstrates the importance of investigating the metallic SLS in order to understand the intrinsic properties of the glass transition. □

Received 18 June; accepted 30 September 1999.

- Ediger, M. D., Angell, C. A. & Nagel, S. R. Supercooled liquids and glasses. *J. Phys. Chem.* **100**, 13200–13212 (1996).
- Grest, G. S. & Cohen, M. H. Liquids, glasses, and the glass transition: a free-volume approach. *Adv. Chem. Phys.* **48**, 454–525 (1981).
- Götze, W. & Sjögren, L. Relaxation processes in supercooled liquids. *Rep. Prog. Phys.* **55**, 241–376 (1992).
- Faupel, F. Diffusion in non-crystalline metallic and organic media. *Phys. Status Solidi A* **134**, 9–59 (1992).

5. Stillinger, F. H. Relaxation and flow mechanisms in "fragile" glass-forming liquids. *J. Chem. Phys.* **89**, 6461–6469 (1988).
6. Kivelson, S. A., Zhao, X., Kivelson, D., Fischer, T. M. & Knobler, C. M. Frustration-limited clusters in liquids. *J. Chem. Phys.* **101**, 2391–2397 (1994).
7. Bernal, J. D. Geometry of the structure of monatomic liquids. *Nature* **185**, 68–70 (1960).
8. Peker, A. & Johnson, W. L. A highly processable metallic glass: $Zr_{41.2}Ti_{13.8}Cu_{12.3}Ni_{10}Be_{22.5}$. *Appl. Phys. Lett.* **63**, 2342–2344 (1993).
9. Bakke, E., Busch, R. & Johnson, W. L. The viscosity of the $Zr_{46.75}Ti_{8.25}Cu_{7.5}Ni_{10}Be_{27.5}$ bulk metallic glass forming alloy in the supercooled liquid. *Appl. Phys. Lett.* **67**, 3260–3262 (1995).
10. Busch, R. & Johnson, W. L. The kinetic glass transition of the $Zr_{46.75}Ti_{8.25}Cu_{7.5}Ni_{10}Be_{27.5}$ bulk metallic glass former-supercooled liquids on a long time scale. *Appl. Phys. Lett.* **72**, 2695–2697 (1998).
11. Inoue, A., Zhang, T. & Takeuchi, A. Ferrous and nonferrous bulk amorphous alloys. *Mater. Sci. Forum* **269–272**, 855–864 (1998).
12. Greer, A. L. Metallic glasses. *Science* **267**, 1947–1953 (1995).
13. Geyer, U. *et al.* Atomic diffusion in the supercooled liquid and glassy states of the alloy $Zr_{41.2}Ti_{13.8}Cu_{12.3}Ni_{10}Be_{22.5}$. *Phys. Rev. Lett.* **75**, 2364–2367 (1995).
14. Geyer, U. *et al.* Small atom diffusion and breakdown of the Stokes–Einstein relation in the supercooled liquid state of the $Zr_{46.75}Ti_{8.25}Cu_{7.5}Ni_{10}Be_{27.5}$ alloy. *Appl. Phys. Lett.* **69**, 2492–2494 (1996).
15. Knorr, K., Macht, M. -P., Freitag, K. & Mehrer, H. Self-diffusion in the amorphous and supercooled liquid state of the bulk metallic glass $Zr_{46.75}Ti_{8.25}Cu_{7.5}Ni_{10}Be_{27.5}$. *J. Non-Cryst. Solids* **250–252**, 669–673 (1999).
16. Fielitz, P., Macht, M. -P., Naundorf, V. & Froberg, G. Diffusion in $ZrTiCuNiBe$ bulk glasses at temperatures around the glass transition. *J. Non-Cryst. Solids* **250–252**, 674–678 (1999).
17. Ehmler, H., Heesemann, A., Rätzke, K., Faupel, F. & Geyer, U. Mass dependence of diffusion in a supercooled metallic melt. *Phys. Rev. Lett.* **80**, 4919–4922 (1998).
18. Tang, X. -P. & Wu, Y. Alignment echo of spin-3/2 9Be nuclei: detection of ultraslow motion. *J. Magn. Reson.* **133**, 155–165 (1998).
19. Tang, X. -P., Busch, R., Johnson, W. L. & Wu, Y. Slow atomic motion in $Zr-Ti-Cu-Ni-Be$ metallic glasses studied by NMR. *Phys. Rev. Lett.* **81**, 5358–5361 (1998).
20. Spiess, H. W. Deuteron spin alignment: a probe for studying ultraslow motions in solids and solid polymers. *J. Chem. Phys.* **72**, 6755–6762 (1980).
21. Fuara, F., Wefing, S. & Kuhs, W. E. Direct observation of tetrahedral hydrogen jumps in ice Ih. *J. Chem. Phys.* **88**, 6801–6809 (1988).
22. Ehlich, D. & Silescu, H. Tracer diffusion at the glass-transition. *Macromolecules* **23**, 1600–1610 (1990).

Acknowledgements

This work was supported by the US Army Research Office, the US National Science Foundation, and the US Department of Energy.

Correspondence and requests for materials should be addressed to Y.W. (yuwewu@physics.unc.edu)

Diamond formation by thermal activation of graphite

A. V. Palnichenko^{*†}, A. M. Jonas[‡], J.-C. Charlier^{*}, A. S. Aronin[†] & J.-P. Issi^{*}

^{*} *Unité de Physico-Chimie et de Physique des Matériaux, Université Catholique de Louvain, Place Croix du Sud 1, B-1348 Louvain-la-Neuve, Belgium*

[†] *Institute of Solid State Physics, Russian Academy of Sciences, Moscow district, 142432, Chernogolovka, Russia*

[‡] *Unité de Chimie et de Physique des Hauts Polymères, Université Catholique de Louvain, Place Croix du Sud 1, B-1348 Louvain-la-Neuve, Belgium*

Synthetic diamond is used in applications ranging from abrasives, tool coatings, bearing surfaces, microelectronics and optics to corrosion protection^{1,2}. The first artificial synthesis used high-pressure techniques to produce diamond as the thermodynamically stable form³, but it can also be grown at low pressures as a metastable carbon phase^{1,2}. Here we report the production of high-purity cubic diamond microparticles (10–100 μm), which form in a highly concentrated carbon-vapour phase, followed by deposition of the crystals on the substrate. The carbon-vapour phase is generated by thermal activation of graphite, and the fast initial growth-rates of diamond, in the range 100–500 $\mu\text{m s}^{-1}$, are at least two orders of magnitude higher than previously reported^{1,2}. We expect that tuning of experimental parameters to optimize the density of the carbon-vapour phase will allow us to grow larger diamond crystals, thereby opening a wider range of potential practical applications.

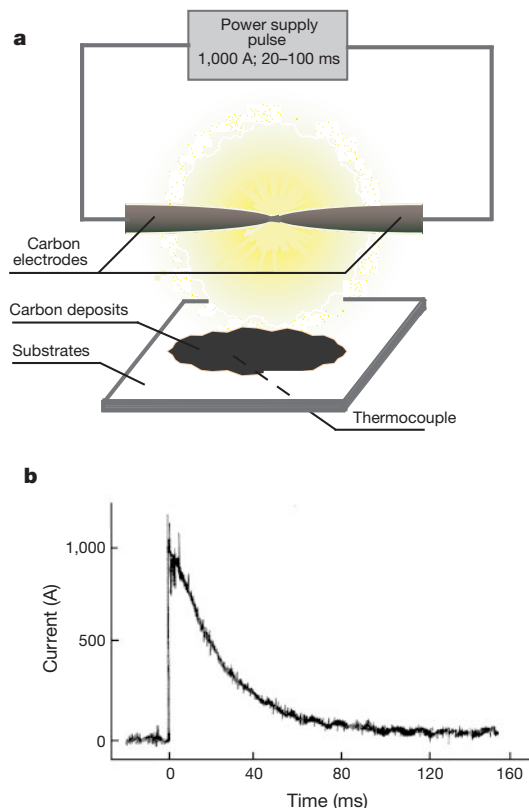


Figure 1 Experimental set-up and duration of the electric pulse discharge for the foundation of diamond microcrystals. **a**, Schematic representation of the thermal activation arc-discharge technique used to synthesize diamond microcrystals. **b**, Typical electric current pulse versus time. During the discharge of the capacitors, the current decays exponentially.

Here we report on the synthesis of high-purity diamond microparticles using a non-equilibrium low-pressure process. Our process is based on a short and intense heat pulse which generates nucleation and growth of diamond in the dense carbon vapour. The carbon plasma is formed in a stainless-steel vacuum-tight chamber by an electric pulse discharge (~ 800 – $1,000$ A/20–100 ms) between two sharpened tips (~ 0.5 mm in diameter) of high purity (99.999%) graphitic electrodes (Fig. 1a). The pulsed current (Fig. 1b) is generated by an electric capacitor (0.5 F) initially charged to 30 V. In order to remove the gaseous impurities before the pulse generation, the chamber is filled with helium, which is then pumped out. This procedure is repeated several times, while a continuous heat treatment of the carbon electrodes at 600–800 K is maintained by means of an electrical current. Dynamic pumping down to 0.01–0.05 torr is carried out through low temperature (77 K) adsorption and standard room-temperature filters connected in series. A set of different substrates, such as nickel, silicon, aluminium, copper and quartz, in the form of 10-cm² plates, are used for the sample deposition, and located at a distance of 5–10 mm from the carbon-electrode tips, in the region of high carbon concentration. During the electric discharge, the substrate is maintained at room temperature, within ~ 10 K, by thermally anchoring it to a heat sink, a ~ 500 g copper block at room temperature. The temperature of the substrate is constantly monitored by a chromel–alumel thermocouple (± 1 K accuracy), with its junction anchored to a region of the substrate just below the carbon tips area. In our experiment, during the electric discharge, the energy density fed to the carbon tips is calculated to be of the order of $\sim 10^7$ W cm⁻².

At the end of the process, hundreds of brilliant colourless particles (~ 10 μm in size) are detectable on the substrate by an

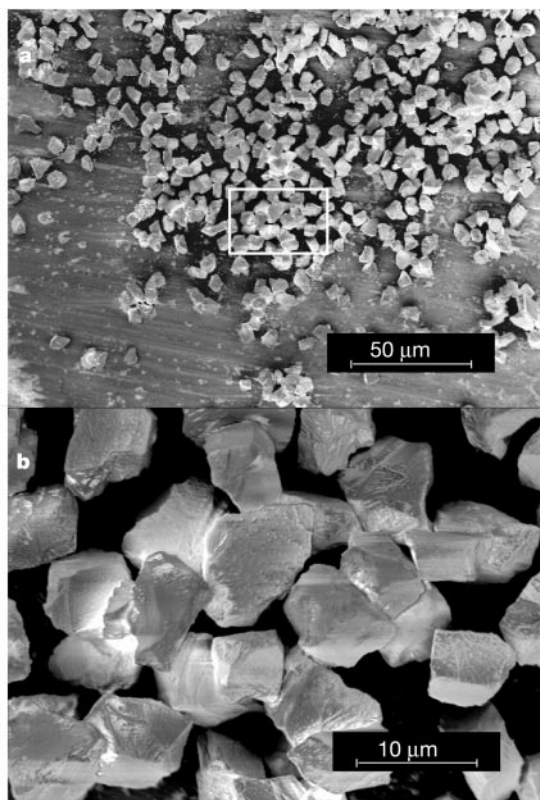


Figure 2 Scanning electron micrographs of diamond microparticles synthesized within a dense carbon-vapour process. A locally high concentration of diamond microparticles is observed on an aluminium substrate in **a**. The average size of these diamond

microcrystallites is $\sim 10 \mu\text{m}$ (area of rectangle is shown enlarged in **b**). We note the random stacking of the diamond microparticles which were not grown on the substrate.

optical polarization microscope. Scanning electron micrographs of these microparticles are shown in Fig. 2. Some larger particles, reaching nearly $100 \mu\text{m}$, have also frequently been observed. The average density of particles on the substrate surface ($\sim 10 \text{ cm}^2$), generated by a single pulse, was determined to be $\sim 10\text{--}50$ microparticles per square centimetre. These microparticles are not distributed randomly, but instead are assembled into dense clusters at various locations on the substrate. The size and the density of these microparticles are found to be independent of the nature of the substrate.

A powder of these microparticles was obtained by gently shaking the samples in hexane, followed by hexane vaporization. The ease with which the powder can be removed from its substrate indicates that the microparticles do not stick to the substrate; a few milligrams of the material could be collected using this simple process. X-ray scattering experiments were conducted on powder samples from about 15 substrates (Fig. 3). The signal corresponds unambiguously to cubic diamond and does not show significant traces of impurities. The width of the reflections is limited by the instrumental resolution, confirming the size and the high degree of perfection of the diamond microparticles. A Fourier-transform infrared (FTIR) spectrum of the diamond microparticles is shown in Fig. 4. The absorption bands at $1,800\text{--}2,500 \text{ cm}^{-1}$ are the well-known signature of multi-phonon absorption processes in diamond⁴. The two weak absorption bands at $1,100 \text{ cm}^{-1}$ and $1,290 \text{ cm}^{-1}$ are attributed to structural defects in our diamond microcrystals, such as small mixed $sp^2\text{--}sp^3$ regions^{5,6}. Nevertheless, the concentration of such structural defects or possible contamination by substitutional nitrogen must be small, otherwise the single-phonon mode (below $1,400 \text{ cm}^{-1}$) would have been infrared-active^{7,8}. This is clearly not the case for the samples investigated here. No absorption bands associated with C–H vibrations ($650\text{--}1,650 \text{ cm}^{-1}$ bending and $2,700\text{--}3,310 \text{ cm}^{-1}$ stretch-

ing) were detected in the infrared spectrum. The absence of impurities in the carbon deposit was also checked using energy dispersive X-ray microanalysis. Therefore, since multi-phonon absorption is dominant, and no absorption from bonds containing nitrogen or hydrogen is observed in the infrared spectrum, the microparticles are assumed to be type Ib or IIa diamond, according to the usual classification⁹.

The Raman spectrum of the microparticles exhibits a sharp first-order peak at $1,331 \text{ cm}^{-1}$ (Fig. 5), which is the characteristic signature of the Raman-active phonon at the Γ point of the Brillouin zone for cubic diamond structures². Both the full-width at half-maximum (FWHM) value of the dominant peak (3.5 cm^{-1}

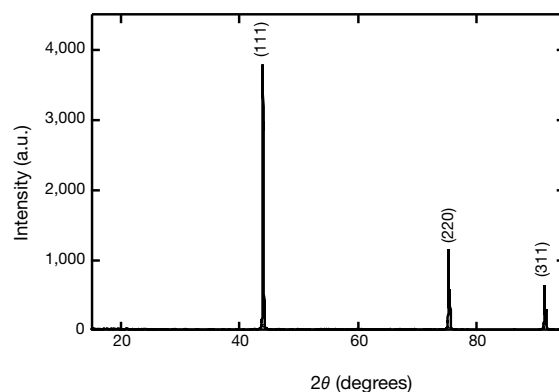


Figure 3 X-ray scattering data acquired on a Siemens D5000 diffractometer in Bragg–Brentano geometry for a large amount of diamond microparticles deposited on a silicon wafer. The reflection peaks, labelled 111, 220 and 311, correspond unambiguously to cubic diamond ($a_0 = 3.567 \text{ \AA}$). The peak splitting is due to the polychromatic nature of the incident beam (Cu $K\alpha 1$ and Cu $K\alpha 2$).

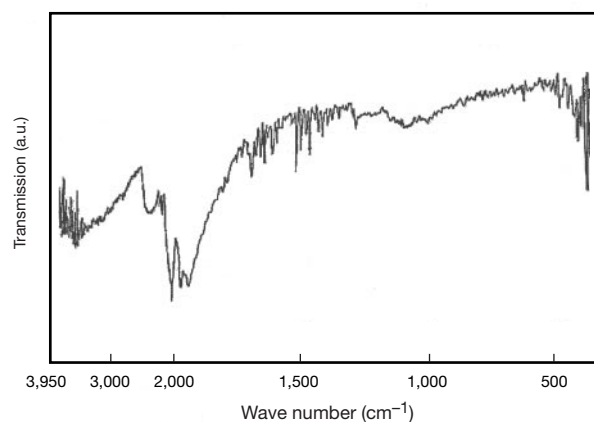


Figure 4 Infrared microscopy spectrum of diamond microparticles. The absorption bands at 1,800–2,500 cm^{-1} are characteristic multiphonon absorption processes in diamond crystals.

width) and the diamond-peak-to-background ratio (of $\sim 200/1$) confirm the crystalline perfection of our vapour-grown diamond. Also, the Raman line FWHM measured for our synthetic diamond is close to that for natural diamonds (1.5–2.5 cm^{-1}), while diamond formed by chemical-vapour deposition (CVD) exhibits a larger value (5–10 cm^{-1}), due to strain and crystal defects in the films¹⁰. The spectrum in Fig. 5 also shows three additional small Raman bands at $\sim 670 \text{ cm}^{-1}$, 1,150 cm^{-1} and 2,100–2,700 cm^{-1} . The origin of the two first peaks is not known, although the peak at 1,150 cm^{-1} was previously reported in the Raman spectra of CVD diamonds^{10,11}. This 1,150 cm^{-1} peak was assumed¹² to originate from the nano-crystalline form of diamond, resulting in a Raman selection rule breakdown (compare with the “D” peak in micro-crystalline graphite). These two peaks could possibly also be attributed to the presence of non-diamond carbon precursors incorporated into the diamond microparticles. The 2,100–2,700 cm^{-1} peak, in the region where scattering by optical phonons is dominant, is the signature of the second-order Raman-active phonon of cubic diamond⁴. The second-order Raman spectrum of our diamond microparticles (inset, Fig. 5) extends to nearly 2,680 cm^{-1} and is very weak below 1,800 cm^{-1} , in agreement with previous studies^{4,13}.

Several identical discharges were also performed, with the substrate located in the low-concentration carbon-gas area, further away (50–100 mm) from the graphitic electrode tips. The average size of the diamond microparticles produced remains $\sim 10 \mu\text{m}$. This result confirms that the diamond particles form in the high-concentration carbon-vapour phase, and not on the substrate surface, as is the case in usual vapour-deposition processes. When the dense carbon-vapour phase is no longer sustained, the completely formed diamond microparticles drop onto the substrate; this explains the random stacking in the aggregates observed in Fig. 2. In such a synthetic process, the diamond microparticles are probably electrostatically charged, which explains the observed self-assembly in the aggregates.

In addition, several runs of sequential electric discharges (up to 500), performed before removing the substrate from the chamber, do not modify either the size or the shapes of the diamond particles. However, the content of the particles on the substrate increases proportionally with the number of pulses in the run. This method may thus be of practical use in the production of relatively large quantities of diamond microcrystals.

Because the carbon gas is generated by an electric current pulse which decays exponentially with time (Fig. 1b), the duration of the pulse is determined to be the time interval from when the current is first switched on to when it has died down to a few amperes. Depending on the electric contact between the carbon tips, the

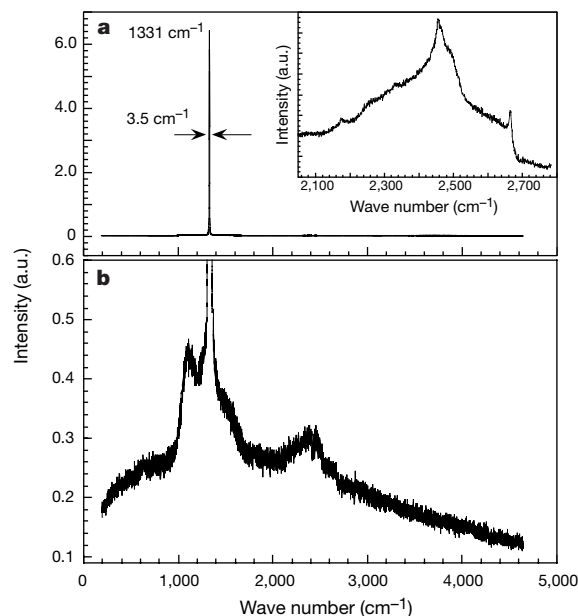


Figure 5 Raman spectra of diamond microcrystals, using an excitation wavelength of 632.8 nm. **a**, The sharp Raman peak for cubic diamond at $1331 \pm 1 \text{ cm}^{-1}$, with a peak line width of 3.5 cm^{-1} , indicates high structural perfection. **b**, The enlarged background spectrum illustrates weaker Raman features. Inset, the second-order Raman spectrum.

duration of the pulse varies within 20–100 ms. As the diamond particles of $\sim 10 \mu\text{m}$ in size are nucleated and grown during the pulse, the corresponding growth rates, over such timescales, are calculated to be $500 \mu\text{m s}^{-1}$ – $100 \mu\text{m s}^{-1}$. As far as we know, such high growth rates have never been reported before for artificial diamond. In addition, a very high current ($>500 \text{ A}$) is only available during the first 10 ms of the discharge (as shown in Fig. 1b), a condition which is certainly required to favour diamond formation. These growth rates are thus minimum values, which could probably be multiplied by a factor of two.

For our synthesis technique, the selective ‘etching’ of competing non-diamond carbon forms by the introduction of atomic gaseous agents in the chamber, such as hydrogen or oxygen, is no longer essential, contrary to many CVD processes^{1,14–16}. The introduction of pure hydrogen gas in the chamber at a pressure of up to 700 torr does not modify either the sizes or the shapes of the diamond particles produced. This suggests that the microscopic mechanisms involved in our diamond formation are probably not the same as those governing more usual chemical vapour deposition techniques. □

Received 13 May; accepted 17 September 1999.

- Angus, J. C. & Hayman, C. C. Low pressure, metastable growth of diamond and “diamondlike” phases. *Science* **241**, 913–921 (1988).
- Yarbrough, W. A. & Messier, R. Current issues and problems in the chemical vapor deposition of diamond. *Science* **247**, 688–696 (1990).
- Bundy, F. P., Hall, H. T., Strong, H. M. & Wentorf, R. H. Man-made diamonds. *Nature* **176**, 51–55 (1955).
- Solin, S. A. & Ramdas, A. K. Raman spectrum of diamond. *Phys. Rev. B* **1**, 1687–1698 (1970).
- Bourdon, E. B. D., Duley, W. W., Jones, A. P. & Prince, R. H. in *Diamond and Diamond-like Films and Coatings* (eds Clausing, R. E. et al.) 297–305 (Plenum, New York, 1991).
- De Vita, A., Galli, G., Canning, A. & Car, R. A microscopic model for surface-induced diamond-to-graphite transitions. *Nature* **379**, 523–526 (1996).
- Sutherland, G. B. B. M., Blackwell, D. E. & Simeral, W. G. The problem of the two types of diamond. *Nature* **174**, 901–904 (1954).
- Kaiser, W. & Bond, W. Nitrogen, a major impurity in common type I diamond. *Phys. Rev. Lett.* **115**, 857–863 (1959).
- Field, J. E. in *Diamond and Diamond-like Films and Coatings* (eds Clausing, R. E. et al.) 17–35 (Plenum, New York, 1991).
- Bachmann, P. K. & Wiechert, D. U. in *Diamond and Diamond-like Films and Coatings* (eds Clausing, R. E. et al.) 677–713 (Plenum, New York, 1991).
- Nemanich, R. J., Glass, J. T., Lucovsky, G. & Shroder, R. E. Raman scattering characterization of carbon bonding in diamond and diamondlike thin film. *J. Vac. Sci. Technol. A* **6**, 1783–1787 (1988).
- Knight, D. S. & White, W. B. Characterisation of diamond films by Raman Spectroscopy. *J. Mater. Res.* **4**, 385–393 (1989).

13. Clark, C. D., Mitchell, E. W. J. & Parsons, B. J. in *The Properties of Diamond* (ed. Field, J. E.) 23–77 (Academic, London, 1979).
14. Eversole, W. G. *Synthesis of Diamond* (US Patents 3030187 & 3030188, 1962).
15. Spitsyn, B. V., Bouilov, L. L. & Derjaguin, B. V. Vapour growth of diamond and other surfaces. *J. Cryst. Growth* **52**, 219–226 (1981).
16. Yoshimoto, M. *et al.* Epitaxial diamond growth on sapphire in an oxidizing environment. *Nature* **399**, 340–342 (1999).

Acknowledgements

A.P. would like to thank S. Tanuma and J. Devaux for discussions, S. Bebelman and X. Vanden Eynde for support in spectrometry techniques. We thank M. S. Dresselhaus for critical reading of the manuscript, and G. Mariage and L. Malcorps for technical help. This work was partially supported through the Belgian Program on Interuniversity Attraction Poles on Reduced Dimensionality Systems, initiated by the Belgian Federal SSTC, and through the Russian Ministry of Scientific and Technical Policy in the frame of the Scientific Program Fullerenes and Atomic Clusters (Carbolite). J.C.C. acknowledges the financial support of the National Foundation for Scientific Research of Belgium.

Correspondence and requests for materials should be addressed to J.-C.C. (e-mail: charlier@pcpm.ucl.ac.be).

Modulation of cadmium uptake in phytoplankton by seawater CO₂ concentration

Jay T. Cullen*, Todd W. Lane†, François M. M. Morel† & Robert M. Sherrell*

* Institute of Marine and Coastal Sciences, Rutgers University, 71 Dudley Road, New Brunswick, New Jersey 08901-8521, USA

† Department of Geosciences, Guyot Hall, Princeton University, New Jersey 08544, USA

The vertical distribution of cadmium in the ocean is characteristic of an algal nutrient^{1,2}, although an underlying physiological basis remains undiscovered. The strong correlation between dissolved cadmium and phosphorus concentrations in sea water has nevertheless been exploited for reconstructing past nutrient distributions in the ocean^{3–5}. In culture experiments, the addition of cadmium accelerates the growth of some marine phytoplankton^{6–9} and increases the activity of carbonic anhydrase, normally a zinc-based metalloenzyme that is involved in inorganic carbon acquisition^{7,9}. Here we show that the concentration of a Cd-carbonic-anhydrase—distinct from Zn-carbonic-anhydrases—in a marine diatom is regulated by the CO₂ partial pressure (p_{CO_2}) as well as by the zinc concentration. Field studies in intensely productive coastal waters off central California demonstrate that cadmium content in natural phytoplankton populations similarly increases as surface-water p_{CO_2} decreases. Incubation experiments confirm that cadmium uptake by natural phytoplankton is inversely related to seawater p_{CO_2} and zinc concentration. We thus propose that biological removal of cadmium from ocean surface waters is related to its utilization in carbonic anhydrase, and is regulated by dissolved CO₂ and zinc concentrations. The dissolved seawater Cd/P ratio would therefore vary with atmospheric p_{CO_2} , complicating the use of cadmium as a tracer of past nutrient concentrations in the upper ocean.

Previous work with laboratory cultures of several species of marine phytoplankton grown at low dissolved inorganic Zn concentrations has shown that Cd additions have a beneficial effect on growth rate, and that this effect is linked to an increase in carbonic anhydrase (CA) activity in the cells^{6–9}. Further, at low inorganic Cd and Zn concentrations (Cd' = 1 pM and Zn' = 3 pM, where M' indicates the sum of inorganic species) relevant to oceanic surface water^{10,11}, the uptake of Cd is increased with decreasing Zn (the usual metal cofactor in CA) and appears to be effected by a separate

transport system¹². CA, which catalyses the formation of CO₂ from HCO₃[–], is part of the inorganic carbon acquisition system whose activity is enhanced at low p_{CO_2} (ref. 7). We have now demonstrated that the diatom *Thalassiosira weissflogii* synthesizes a Cd-CA distinct from the other cellular Zn-CA (T.W.L. and F.M.M., unpublished results) as is illustrated in the native protein gels of Fig. 1 (lanes 6–9) showing distinct bands of ¹⁰⁹Cd and ⁶⁵Zn coeluting with CA activity. The Cd-CA is present in the cells even at an inorganic Zn concentration that allows maximum growth rate (Zn' = 15 pM; lanes 4–6), and its expression is regulated by p_{CO_2} as well as Zn' (lanes 1–6). The relative amounts of ¹⁰⁹Cd present indicate that the concentration of Cd associated with the enzyme increases more than ten times when p_{CO_2} is decreased from 750 to 100 p.p.m. (lanes 1 and 3), and by about 50% when Zn' is decreased from 15 to 3 pM (lanes 5 and 2).

As the phytoplankton Cd-CA is the only known biological function for Cd, its uptake by the biota in surface sea water may result from its use in such enzymes. We would then expect that Cd uptake by phytoplankton in the oceans should be modulated by ambient p_{CO_2} and Zn. We measured the Cd content of natural phytoplankton assemblages along environmental p_{CO_2} gradients during the April 1997 upwelling period within 50 km of the central California coast between Monterey Bay and Santa Barbara Channel (>0.45 μm particulate Cd/P ratios were determined directly by HR-ICP-MS (high resolution inductively coupled plasma mass spectrometry) C/P = 106–120, $n = 4$, for particulate samples from Monterey Bay, indicating that living biomass dominates particle composition). Surface-water p_{CO_2} varied from 250 p.p.m. in regional plankton blooms, found both in Monterey Bay and in the central Santa Barbara Channel, to >800 p.p.m. in recently upwelled water off Big Sur. The phytoplankton Cd/P ratio was inversely correlated to surface-water p_{CO_2} , with values for the >0.45-μm biomass varying from 0.06 to 0.7 mmol mol^{–1} over the sampled range of p_{CO_2} (Fig. 2). The lowest value, found in recently upwelled water, was six times lower than the dissolved Cd/P ratio in the upwelled water, while the highest—observed in regional high-chlorophyll blooms at p_{CO_2} = 250–300 p.p.m.—was twice the upwelled Cd/P ratio.

We also determined the Zn/P ratio of our samples; cellular Cd and Zn concentrations were inversely correlated, particularly in the large

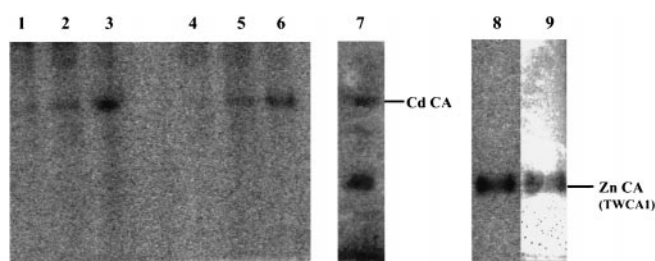


Figure 1 Modulation of Cd-carbonic-anhydrase (Cd-CA) by p_{CO_2} and Zn. Lanes 1–6 are phosphorimages of 10% native polyacrylamide gels of diatom proteins labelled with ¹⁰⁹Cd and grown under different concentrations of Zn and p_{CO_2} . Lanes 1–3, 3 pM Zn', 30 pM Cd'; lanes 4–6, 15 pM Zn', 30 pM Cd'; lanes 1 and 4, bubbled with 750 p.p.m. CO₂; lanes 2 and 5, 350 p.p.m. CO₂; lanes 3 and 6, 100 p.p.m. CO₂. Lane 7, CA activity assay of lane 6. Lane 8 is a phosphorimage of 10% native gel of diatom proteins labelled with ⁶⁵Zn grown with 15 pM Zn' and 100 p.p.m. CO₂. Lane 9, CA activity assay of lane 8 (TWCA1 is the main Zn-containing isoform of CA in *T. weissflogii*). Under incubation conditions (20 °C, salinity = 30), p_{CO_2} values of 100, 350 and 750 p.p.m. correspond to aqueous concentrations of CO₂ of 3.3, 11.7, and 25.0 μM, respectively. *T. weissflogii* harvested in mid-exponential phase was resuspended in buffer (10 mM Tris-HCl, pH 7.2, 1% Triton X100) and lysed by sonication. Protein representing equal cell numbers was loaded in each lane of the gel. CA activity was detected by the bromothymol blue pH drift assay described elsewhere²⁷. CA activity appears as dark on a lighter background. Relative amounts of radioactivity present in each band was determined by phosphor-imaging.

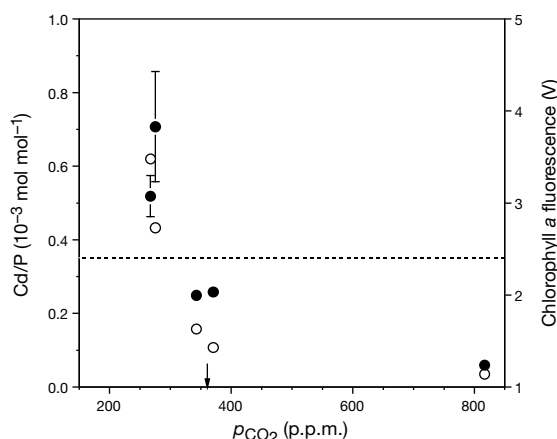


Figure 2 Phytoplankton Cd/P and chlorophyll *a* fluorescence versus p_{CO_2} in field samples. Fluorescence is reported as a relative analogue signal in volts (open circles). Samples were collected at locations within 50 km off central California in April 1997 (p_{CO_2} data courtesy of T. Takahashi, C. Sweeney and A. van Geen). The highest Cd/P ratio ($>0.45 \mu\text{m}$; closed circles) was observed in two bloom areas in Monterey Bay and Santa Barbara Channel. The calculated aqueous concentration of CO_2 at bloom sites reached a minimum of $12 \mu\text{M}$. Error bars (some smaller than symbol) are estimated standard deviation of ratios, largely a function of filter blank reproducibility. Dotted line shows upwelled dissolved Cd/P ratio, and arrow indicates mean atmospheric p_{CO_2} . A novel sampling technique and analytical method was applied to measure particulate metals and P in field samples²⁸. Surface-water samples collected with a trace-metal-clean pumping system were filtered through a series of three filters, ending with acid-leached $0.45\text{-}\mu\text{m}$ polysulphone filters (Supor, Gelman, Ann Arbor, Michigan). Samples were stored at -20°C in the dark until analysed in the laboratory. Phytoplankton were digested in concentrated HNO_3 and HF acids (SeaStar, Sidney, British Columbia, Canada), and analysed for trace metals and P by HR-ICP-MS (Element, Finnigan-MAT, Bremen, Germany) using the method of standard additions. To calculate the composition of biogenic particles, Fe and Al were determined as tracers of inorganic particles to estimate non-biological fractions of metals and P in particulate samples. Based on average crustal ratios, terrigenous clays contributed $<1\%$ of total Cd, Zn and P in all samples discussed. Estimated P associated with authigenic Fe oxide phases was $<5\%$ of total measured P, based on P/Fe in Fe oxyhydroxides precipitated within hydrothermal plumes²⁹. p_{CO_2} was determined by continuous (every 2.5 min) underway equilibrator measurements from $\sim 2\text{ m}$ depth (T. Takahashi and C. Sweeney, personal communication).

size fractions (Fig. 3). The largest size fraction ($>53 \mu\text{m}$), which consisted primarily of large diatoms (mostly *Nitzschia* spp.) also had by far the highest Cd content in the bloom samples: $\text{Cd/P} = 1\text{--}6 \text{ mmol mol}^{-1}$ at $270 \text{ p.p.m. } p_{\text{CO}_2}$, ten times higher than in the same size fraction at $340\text{--}820 \text{ p.p.m.}$ This result is significant as CO_2 uptake by large cells ($>30 \mu\text{m}$) may be diffusion-limited¹³, implying increased requirement for CA at the dissolved CO_2 concentrations seen in our California bloom stations. The size-fractionated data provide evidence against adsorption of Cd on particles (living or dead) as a mechanism for removal, because larger particles with lower surface to volume ratios have higher, not lower, Cd/P ratios than smaller particles. As large, high-Cd diatoms account for a disproportionate fraction of export production^{14–16}, dissolved Cd may be removed preferentially to phosphate in productive, low- p_{CO_2} waters.

Because several chemical and biological parameters covary with p_{CO_2} in the surface waters off California, the observed correlations may not represent a causal dependency of particulate Cd on p_{CO_2} and Zn. To establish more directly the role of p_{CO_2} and Zn concentration in modulating Cd concentrations in marine phytoplankton, we performed short term (5–6 hours) ^{109}Cd and $\text{H}^{14}\text{CO}_3^-$ uptake experiments in shipboard incubations at varying Zn and p_{CO_2} , using the natural assemblage sampled from this region in May 1998. Calculated Cd/P (assuming the Redfield value of

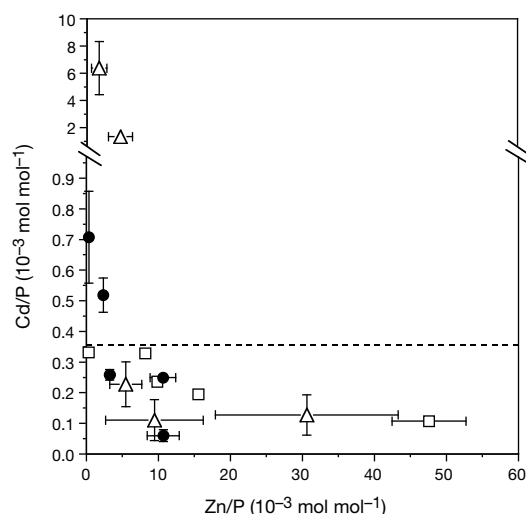


Figure 3 Cd/P versus Zn/P for phytoplankton off central California. Total $>0.45 \mu\text{m}$ biomass (filled circles), $5\text{--}53 \mu\text{m}$ fraction (open squares), and $>53 \mu\text{m}$ fraction (open triangles), for assemblages dominated by large chain-forming diatoms. Note the change in the scale to accommodate the very high Cd content of the largest size fraction in bloom regions.

$\text{C/P} = 106 \text{ mmol mol}^{-1}$) was substantially lower than in 1997 field samples. The reason for this difference is unknown, but may reflect a different flora or inhibition of Cd uptake by relatively high dissolved Mn concentrations in incubation seawater (40 nM , M. Wells, personal communication)¹². This difference notwithstanding, the Cd content of the phytoplankton was found to depend on the concentrations of dissolved CO_2 and Zn. Cells acclimated to grow at $100 \text{ p.p.m. } p_{\text{CO}_2}$ and ambient Zn (0.7 nM ; M. Wells, personal communication) had intracellular Cd concentrations (measured as steady-state Cd/C uptake ratios) twice those of cells maintained at 350 and $800 \text{ p.p.m. } \text{CO}_2$ (Fig. 4). In all cases, we observed a significant decrease ($50 \pm 20\%$) in Cd uptake when 25 nM Zn was added to the samples. Nonetheless, at low p_{CO_2} (100 p.p.m.), the uptake of Cd was two or three times greater than at higher p_{CO_2} , even in the presence of a high Zn concentration.

The incubation experiments confirm that the variations in biological uptake of Cd in waters off California are indeed modulated by p_{CO_2} and Zn. Such modulation is consistent with the observed regulation of Cd-CA in culture, and supports the hypothesis that Cd uptake in surface sea water results from its utilization in Cd-CA in phytoplankton. On this basis, we predict that observed variations in upper-ocean Cd cycling reflect variations in ambient Zn and, particularly, CO_2 concentrations as observed in the waters off central California. For example, we suggest that high Cd uptake and vertical export can explain recent observations of sharply reduced dissolved Cd/P ($0.05\text{--}0.15 \text{ mmol mol}^{-1}$; refs 17–19) in the low- p_{CO_2} ($\sim 100 \text{ p.p.m.}$ undersaturated; ref. 20) subantarctic surface waters, compared to Cd/P ($0.3 \text{ mmol mol}^{-1}$) in Antarctic waters south of the Polar Front^{18,21,22}. Further, production of Cd-rich biomass is consistent with vertical profiles of dissolved Cd versus P north of the Polar Front, which show regeneration of organic matter containing high Cd/P ($\sim 0.5 \text{ mmol mol}^{-1}$; ref. 18).

A modulation of Cd uptake by p_{CO_2} also suggests that the biogeochemical cycle of Cd should respond to major changes in global climate. We postulate that the relationship between Cd and P in modern ocean surface waters is a result of dissolved CO_2 and Zn availability to large phytoplankton that dominate export production. Past changes in CO_2 and/or Zn availability may therefore have altered this relationship either globally or regionally. For example, at the low p_{CO_2} of the glacial atmosphere ($\sim 200 \text{ p.p.m.}$; ref. 23, 24), the p_{CO_2} of the subantarctic surface water (which may have been further

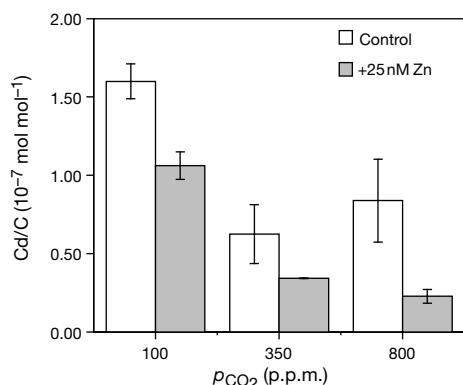


Figure 4 Cd/C uptake of phytoplankton grown at varying p_{CO_2} and $[\text{Zn}(\text{aq.})]$. A natural phytoplankton assemblage from Monterey Bay was acclimated to three different p_{CO_2} levels at ambient (~ 0.7 nM) and amended Zn (+25 nM). Error bars represent propagated standard deviation based on duplicate measurements of Cd and C uptake rates. Surface-water samples (2-l bottles from 20-m depth) were collected in May 1998 and incubated on board ship using trace-metal-clean techniques at ambient sea surface temperature ($13 \pm 2^\circ\text{C}$), bubbled with air– CO_2 mixtures at 100, 350 and 800 p.p.m. p_{CO_2} (Scott Specialty Gases, Plumsteadville, Pennsylvania). Following an acclimation period (20–27 h, ~ 1 doubling), samples were transferred to 250-ml acid-cleaned polycarbonate bottles (Nalgene), Zn was added to amendment bottles, and Cd and C uptake were measured in separate duplicate subsamples (250 ml) by short-term (5–6 h) uptake of added $\text{H}^{14}\text{CO}_3^-$ and carrier-free ^{109}Cd (Zn and radioisotopes pre-equilibrated with $0.2\text{-}\mu\text{m}$ -filtered surface sea water during acclimation time). The total Cd addition to incubation bottles calculated from the specific activity of the primary stock was 10 pM. After gentle (<150 mm Hg) filtration onto $3\text{-}\mu\text{m}$ polycarbonate membrane filters, cells were rinsed with a 4 mM diethylenetriamine penta-acetic acid solution for 5 min, then washed with $0.2\text{-}\mu\text{m}$ -filtered sea water to remove extracellular surface adsorbed ^{109}Cd , and assayed for ^{109}Cd and ^{14}C activity by liquid scintillation counting. The Cd/C uptake ratio was determined using total inorganic carbon calculated for each p_{CO_2} level (at measured $T = 13^\circ\text{C}$, salinity = 33.135, and total alkalinity = $2,282\text{ }\mu\text{equiv. kg}^{-1}$) and dissolved Cd = 0.7 nM throughout. Although C uptake (per chl a) varied with p_{CO_2} for both Zn levels, changes in C uptake were of secondary importance compared to variable Cd uptake in determining the differences observed in Cd/C ratios across treatments. p_{CO_2} levels of 100, 350 and 800 p.p.m. correspond to calculated aqueous CO_2 concentrations of 4, 14 and $32\text{ }\mu\text{M}$, assuming solubility equilibrium.

undersaturated due to increased productivity during the Last Glacial Maximum, ref. 25) would have been <200 p.p.m., causing the preferential removal of Cd to be more intense than it is now. The resulting low dissolved Cd/P ratio in these waters would then have to be taken into account when interpreting Cd/Ca in fossil planktonic foraminifera as a proxy for surface-water phosphate concentrations²⁶. Our results suggest that a region acting as a substantial atmospheric CO_2 sink probably exports Cd-rich organic matter, driving the surface water dissolved Cd/P ratio below the upwelled deep water ratio. If so, evidence for a lower glacial Cd/Ca ratio in planktonic foraminifera cannot be interpreted directly as more efficient surface-water nutrient utilization. □

Received 24 May; accepted 10 September 1999.

- Boyle, E. A., Sclater, F. & Edmond, J. M. On the marine geochemistry of cadmium. *Nature* **263**, 42–44 (1976).
- Bruland, K. W. Oceanographic distributions of cadmium, zinc, nickel, and copper in the north Pacific. *Earth Planet. Sci. Lett.* **47**, 176–198 (1980).
- Boyle, E. A. Cd and ^{13}C paleochemical ocean distributions during the stage 2 glacial maximum. *Annu. Rev. Earth Planet. Sci.* **20**, 245–287 (1992).
- van Geen, A. et al. Evidence from Cd/Ca ratios in foraminifera for greater upwelling of California 40,000 years ago. *Nature* **358**, 54–56 (1992).
- Rosenthal, Y., Boyle, E. A. & Labeyrie, L. Last glacial maximum paleochemistry and deepwater circulation in the Southern Ocean: Evidence from foraminiferal cadmium. *Paleoceanography* **12**, 778–787 (1997).
- Price, N. M. & Morel, F. M. M. Cadmium and cobalt substitution for zinc in a marine diatom. *Nature* **344**, 658–660 (1990).
- Morel, F. M. M. et al. Zinc and carbon colimitation of marine phytoplankton. *Nature* **369**, 740–742 (1994).

- Lee, J. G. & Morel, F. M. M. Replacement of zinc by cadmium in marine phytoplankton. *Mar. Ecol. Prog. Ser.* **127**, 305–309 (1995).
- Lee, J. G., Roberts, S. B. & Morel, F. M. M. Cadmium: a nutrient for the marine diatom *Thalassiosira weissflogii*. *Limnol. Oceanogr.* **40**, 1056–1063 (1995).
- Bruland, K. W. Complexation of cadmium by natural organic ligands in the central North Pacific. *Limnol. Oceanogr.* **37**, 1008–1017 (1992).
- Bruland, K. W. Oceanic zinc speciation: complexation of zinc by natural organic ligands in the central North Pacific. *Limnol. Oceanogr.* **34**, 267–283 (1989).
- Sunda, W. G. & Huntsman, S. A. Control of cadmium concentrations in a coastal diatom by interactions among free ionic Cd, Zn, and Mn in seawater. *Environ. Sci. Technol.* **32**, 2961–2968 (1998).
- Riebesell, U., Wolf-Gladrow, D. A. & Smetacek, V. Carbon dioxide limitation of marine phytoplankton growth rates. *Nature* **361**, 249–251 (1993).
- Goldman, J. C. Potential role of large oceanic diatoms in new primary production. *Deep-Sea Res.* **40**, 159–168 (1993).
- Dugdale, R. C. & Wilkerson, F. P. Silicate regulation of new production in the equatorial Pacific upwelling. *Nature* **391**, 270–273 (1998).
- Smetacek, V. Diatoms and the silicate factor. *Nature* **391**, 224–225 (1998).
- Frew, R. D. & Hunter, K. A. Influence of Southern Ocean waters on the cadmium-phosphate properties of the global ocean. *Nature* **360**, 144–146 (1992).
- Löscher, B. M., de Jong, J. T. M. & de Baar, H. J. W. The distribution and preferential biological uptake of cadmium at 6°W in the Southern Ocean. *Mar. Chem.* **62**, 259–286 (1998).
- Nolting, R. F. & de Baar, H. J. W. Behaviour of nickel, copper, zinc and cadmium in the upper 300 m of a transect in the Southern Ocean ($57^\circ\text{--}62^\circ\text{S}$, 49°W). *Mar. Chem.* **45**, 225–242 (1994).
- Takahashi, T. et al. Global air-sea flux of CO_2 : An estimate based on measurements of sea-air p_{CO_2} difference. *Proc. Natl Acad. Sci.* **94**, 8292–8299 (1997).
- de Baar, H. J. W., Saager, P. M., Nolting, R. F. & van der Meer, J. Cadmium versus phosphate in the world ocean. *Mar. Chem.* **46**, 261–281 (1994).
- Rutgers van der Loeff, M., Helmers, E. & Kattner, G. Continuous transects of cadmium, copper, and aluminium in surface waters of the Atlantic Ocean, 50°N to 50°S : Correspondence and contrast with nutrient-like behaviour. *Geochim. Cosmochim. Acta* **61**, 47–61 (1997).
- Fischer, H., Wahlen, M., Smith, J., Mastroian, D. & Deck, B. Ice core records of atmospheric CO_2 around the last three glacial terminations. *Science* **283**, 1712–1714 (1999).
- Petit, J. R. et al. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **399**, 429–436 (1999).
- Kumar, N. et al. Increased biological productivity and export production in the glacial Southern Ocean. *Nature* **378**, 675–680 (1995).
- Rickaby, R. E. M. & Elderfield, H. Planktonic foraminiferal Cd/Ca: Paleonutrients or paleotemperature? *Paleoceanography* **14**, 293–303 (1999).
- Roberts, S. B., Lane, T. W. & Morel, F. M. M. Carbonic anhydrase in the marine diatom *Thalassiosira weissflogii* (Bacillariophyceae). *J. Phycol.* **33**, 845–850 (1997).
- Cullen, J. T. & Sherrell, R. M. Techniques for determination of trace metals in small samples of size-fractionated particulate matter: phytoplankton metals off central California. *Mar. Chem.* (in the press).
- Feely, R. A. et al. Hydrothermal plume particles and dissolved phosphate over the superfast-spreading southern East Pacific Rise. *Geochim. Cosmochim. Acta* **60**, 2297–2323 (1996).

Acknowledgements

We thank A. van Geen and M. Wells for providing shipboard space and supporting data; T. Takahashi and C. Sweeney for underway p_{CO_2} measurements; Y. Rosenthal, P. Falkowski, I. Berman-Frank and M. Behrenfeld for comments on an earlier version of the manuscript; P. Tortell for assistance at sea; and P. Field and I. Shaperdorth for assistance in the laboratory. This work was supported by the NSF and the DOE.

Influence of NO_x emissions from ships on tropospheric photochemistry and climate

Mark G. Lawrence & Paul J. Crutzen

Max-Planck-Institut für Chemie, Abteilung Luftchemie, PF 3060, 55020 Mainz, Germany

Emissions of nitrogen oxides (NO_x , the sum of NO and NO_2) from fossil-fuel burning dominate the NO_x burden of the lower troposphere in many regions¹. These emissions increase tropospheric ozone and hydroxyl-radical concentrations over their natural 'background' levels, thereby increasing the oxidizing power of the atmosphere². Fossil-fuel emissions of NO_x (refs 3, 4) account for about half of the global NO_x source to the atmosphere; other significant sources are from biomass burning⁵, soil emissions⁶, aircraft exhausts⁷ and lightning⁸, all primarily continental. However,

ocean-going ships burning fossil fuels may also contribute a significant fraction ($>10\%$) to global NO_x production⁹. Here we use NO_x emission data and a high-resolution chemistry–transport model to estimate that ship NO_x emissions result in a more than 100-fold increase in surface NO_x concentrations in heavily traversed ocean regions. This enhancement has a notable effect on modelled surface ozone and hydroxyl-radical concentrations. In particular, a predicted fivefold increase in the July hydroxyl-radical burden over the northern Atlantic and Pacific oceans would be expected to reduce the atmospheric lifetimes of reactive greenhouse gases—such as methane—as well as to increase aerosol production rates and cloud reflectivities, therefore exerting a cooling influence on the climate.

In the GEIA (Global Emissions Inventory Activity) database of NO_x emissions³ produced by the burning of fossil fuels, currently the most widely used emissions database for global models, ships were assumed to produce 1.6 Tg N yr^{-1} as NO_x ($1 \text{ Tg} = 10^{12} \text{ g}$). However, these emissions were nearly all assigned to ports and inland and coastal waterways, and only include a small open ocean component in the northern Atlantic (Fig. 1a). More recent research^{4,9} indicates a considerable open-ocean component to the ship NO_x emissions and estimates the total to be about twice the value used in GEIA. Corbett and colleagues⁹ compute 3.08 (2.66 – 4.00) Tg N yr^{-1} from ships for 1993: this figure was arrived at by combining marine diesel engine emission factors¹⁰ (17 or 27 kg N per t fuel for medium- and slow-speed engines, respectively) with information on registered engine types¹¹ ($\sim 95\%$ of commercial marine vessels use slow- or fast-type diesel engines) and the total global marine fuel consumption¹² (about $150 \times 10^6 \text{ t fuel yr}^{-1}$). Using a similar approach, Olivier and colleagues⁴ have revised their previously lower estimate of $0.23 \text{ Tg N yr}^{-1}$ upwards to 2.6 Tg N yr^{-1} for 1990, or 2.9 Tg N yr^{-1} for 1995, with an estimated uncertainty of 30% (J. Olivier, personal communication).

To assess the importance of ship NO_x emissions, we use the global tropospheric chemistry–transport model MATCH-MPIC^{13,14}, described in the Methods section. Two model runs are considered, one without NO_x emissions from ships ('SHIPS0'), and one with 3 Tg N yr^{-1} production by ships ('SHIPS3'). We use the spatial distribution of fossil-fuel NO_x emissions from EDGAR⁴ (Environmental Database for Global Atmospheric Research), depicted in Fig. 1b. The distribution of ship emissions in EDGAR was derived from the world's main shipping routes and their traffic intensity^{15,16}. Our simulations are at a relatively high horizontal resolution of about $2^\circ \times 2^\circ$ (192×96 grid cells). For the present study we neglect the oxidation of non-methane hydrocarbons, as their oxidation schemes are computationally prohibitive at this high horizontal resolution, and because they are expected to mainly affect continental regions¹⁷. We also neglect potential O_3 loss induced by reactions of NO_3 and N_2O_5 on seasalt particles¹⁸, which may need to be considered in future assessments.

The global surface distribution of NO_x as computed by the SHIPS3 model run is shown for July in Fig. 1c. NO_x mixing ratios are greatest over the continents, reflecting the source distribution. They are also relatively high ($>100 \text{ pmol mol}^{-1}$) near the main shipping routes, contrasting with the much lower levels ($<10 \text{ pmol mol}^{-1}$) which generally occur over oceanic regions. Relative to the SHIPS0 run (Fig. 1d), the inclusion of ship NO_x emissions results in an increase in surface concentrations of NO_x of more than a factor of two throughout most of the northern Atlantic, Pacific and Indian oceans, exceeding two orders of magnitude in the shipping lanes. The highest ratios are found in regions which start with very low NO_x levels (of the order of 1 pmol mol^{-1}) in the SHIPS0 run; however, it is at low NO_x levels that NO_x additions to the atmosphere have the largest relative effect on concentrations of ozone and hydroxyl radicals¹⁹.

Enhanced levels of NO_x were measured in the marine boundary layer (MBL) over the North Atlantic during the OCTA-2 campaign

(August 1993). Observations^{20,21} between the surface and 1 km altitude showed a median (average) NO mixing ratio of 46.3 (45.0) pmol mol^{-1} for flights in a region from 40°N to 55°N and 20°W to 10°W . Average daytime NO levels from the SHIPS0 and SHIPS3 model runs for this region are 11.7 and $36.2 \text{ pmol mol}^{-1}$, respectively. Similarly, in the region from 40°N – 50°N and 70°W – 55°W , observations found the median (average) NO to be 27.3 (30.1) pmol mol^{-1} , while the model runs computed 21.3 and $29.2 \text{ pmol mol}^{-1}$, respectively. On the other hand, there are also

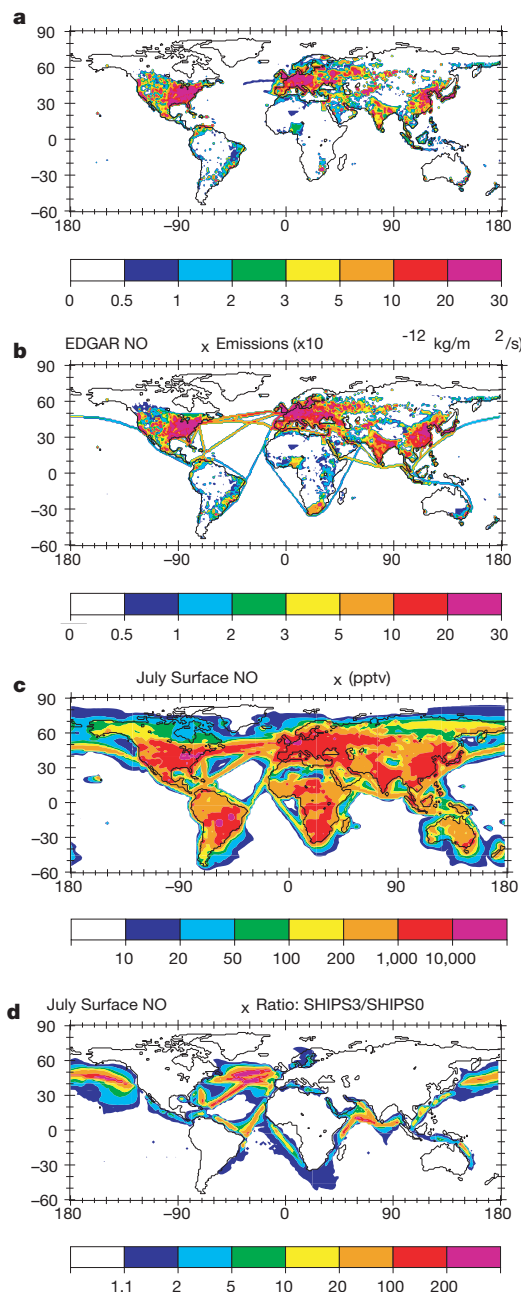


Figure 1 Fossil-fuel NO_x emissions and surface NO_x levels. **a**, Fossil-fuel NO_x emissions ($10^{-12} \text{ kg m}^{-2}$) from GEIA³, used in our previous simulations; **b**, fossil-fuel NO_x emissions ($10^{-12} \text{ kg m}^{-2}$) from EDGAR⁴, with the ship NO_x emissions scaled to total 3 Tg N yr^{-1} globally, used in the SHIPS3 simulation (for the SHIPS0 run the EDGAR emissions are also used, except ship emissions are set to zero; for both runs, the emissions are assumed constant throughout the year); **c**, the July surface NO_x distribution (pmol mol^{-1}) from the SHIPS3 model run; **d**, the ratio of July surface NO_x in the SHIPS3 run versus the SHIPS0 run (a similar result, though with slightly lower ratios, is obtained for January, not shown).

observations of much lower NO levels in some regions of the North Atlantic MBL. Unfortunately, the observations available at present are far too sparse to provide a clear picture of the actual average MBL NO_x levels in the shipping lanes.

With NO_x emissions from ocean-going ships of the magnitude suggested by Corbett and colleagues⁹, the repercussions for the concentrations of other trace gases, in particular O₃ and the hydroxyl radical, OH[•], are significant. Figure 2a, b shows the ratios of July surface O₃ and OH[•] from the SHIPS3 run relative to the

SHIPS0 run. The concentration of O₃ is computed to increase by more than a factor of 2 over the central North Atlantic and Pacific. Even more notable is the effect on OH[•], which increases by more than 20% for much of the northern oceans, and more than a factor of five in the shipping lanes away from the coasts. The computed large increase extends throughout the MBL (Fig. 2c), and yields surface OH[•] concentrations in the SHIPS3 run (Fig. 2d) in the northern Atlantic which are as high as typically found over the continents (in strong contrast to the SHIPS0 run, not shown, in which MBL OH[•] concentrations do not exceed 2×10^6 molecules cm⁻³). This in turn affects the oxidation of tropospheric gases. For instance, in the SHIPS3 run, the breakdown rate of CH₄ in the Atlantic MBL from 30° N–60° N is 64% greater than in the SHIPS0 run. For the entire Northern Hemisphere MBL, the CH₄ oxidation rate is increased by 19% due to ship NO_x emissions. This shift towards a greater relative oxidizing capacity of the atmosphere over the oceans, in particular the northern Atlantic and Pacific, may also be enhanced by the finding¹⁷ that non-methane hydrocarbons (not included in our model run) tend to significantly reduce OH[•] concentrations over the forested continents. From a global perspective, the reduction in the computed CH₄ lifetime from 9.6 yr in the SHIPS0 run to 9.2 yr in the SHIPS3 run brings it closer to the current observation-based estimates of 8–9 years (refs 22, 23).

The computed enhancement of atmospheric OH[•] concentrations will also affect aerosol production. A primary precursor for maritime non-seasalt-sulphate aerosol is dimethyl sulphide which is released from the ocean waters and oxidized in the troposphere. In the shipping lanes of the North Atlantic during summer, the lifetime of dimethyl sulphide molecules before oxidation by OH[•] is reduced from about 36 hours to 9 hours (assuming 1×10^6 and 4×10^6 molecules of OH[•] per cm³ in the SHIPS0 and SHIPS3 runs, respectively, using the rate constant²⁴ for 15 °C). As the mixing time for MBL air with the overlying free troposphere is of the order of a few days, this change in the lifetime can make a large difference in how much dimethyl sulphide can be transported out of the MBL before it is oxidized, which has been found to be an important parameter in an aerosol formation model²⁵.

The formation of new particles also depends strongly on the gas-phase oxidation of SO₂ to H₂SO₄ via reaction with OH[•]; this reaction competes with aqueous-phase oxidation of SO₂ (by H₂O₂ and O₃), which only adds to already existing particle masses. Increased OH[•] concentrations reduce the SO₂ lifetime in or near the North Atlantic shipping lanes in the summertime from about 12 to 3 days, leading to a greater fractional oxidation of SO₂ by OH[•], and thus to an increased likelihood of forming new particles there. This in turn results in enhanced reflection of solar radiation, both directly by aerosols and indirectly via increased cloud-droplet number densities, thereby causing a cooling of the region. Capaldo and colleagues²⁶ have recently considered a similar argument for the effects of SO₂ emissions from ships and, without invoking an increase in OH[•] concentration, also estimate a notable cooling. As the overall climate effect of ship emissions is likely to be considerably larger than that due to either NO_x or SO₂ ship emissions alone, particularly given the co-location of the emissions, future studies with coupled models are needed for a more accurate assessment of the issue.

The two largest uncertainties in this study are the distribution and emission factor of the ship NO_x emissions. The shipping routes from EDGAR (Fig. 1b) are probably too narrow; an alternative approach using positions at the times of ship reports and log entries⁹ finds the main emissions to occur in the same regions, though more spread out. As photochemical production of O₃ per added NO_x molecule is more efficient at lower NO_x concentrations¹⁹, more diffuse emissions would result in an even greater effect on O₃ than computed in this study, and in a broader region of significantly enhanced OH[•] levels.

The ship NO_x emission factor assumed here¹⁰ was derived from

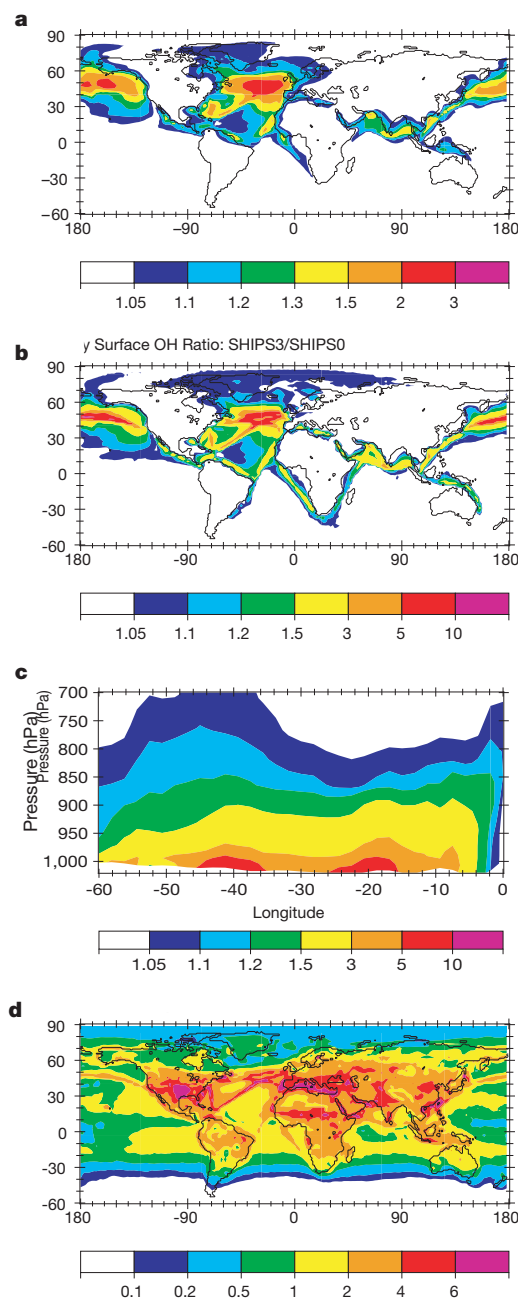


Figure 2 Effects of ship NO_x emissions on O₃ and OH[•]. Depicted are ratios of July surface-level O₃ (**a**) and OH[•] (**b**) from the SHIPS3 run relative to the SHIPS0 run; **c**, a longitude-height slice of the OH[•] ratio (SHIPS3/SHIPS0) over the North Atlantic along 45° N; and **d**, the surface OH[•] distribution ($\times 10^6$ molecules cm⁻³) from the SHIPS3 run (24-hour averages), in which the oceanic shipping lanes are evident. In January (not shown), due to less photochemical activity in the Northern Hemisphere, O₃ only increases significantly in the Southern Hemisphere tropical segments of the main shipping routes, while OH[•] increases in the same regions are shown in **b**, although to a lesser extent (a factor of 2–3 increase in the central northern oceans).

measurements on about 50 ships at various steady-state engine loads. A recent analysis²⁷ of *in situ* aircraft measurements of NO_x in the exhaust plumes of four slow-speed diesel ships indicates a range of 8.4–14 kg N per t fuel, about 0.33–0.5 times the value assumed above. It is unknown whether the discrepancy is due to problems with the techniques of NO_x emission measurement, or whether it is because only a very few ships could be considered in the latter study. Furthermore, it is uncertain how representative either of these emission factors can be for the current global fleet of over 100,000 registered vessels. Resolving these issues is critical for providing an accurate base of information for future policy decisions. In particular, extensive measurements in and around the main shipping routes are needed to obtain a picture of the influence of the ship NO_x source on the oceanic NO_x distribution and its further effect on O₃ and OH[•] concentrations.

The effects of ocean-going ship emission of NO_x on the tropospheric oxidizing power that we calculate, and the potential for cooling effects, are expected to increase notably over the coming decades, as ship traffic is currently growing by about 3% yr⁻¹ globally⁹, and even more strongly in the tropics. The CO₂ release to the atmosphere from ships (120–150 Tg C yr⁻¹; ref. 9) is about 2–2.5% of the global release of CO₂ from fossil-fuel burning²⁸. In comparison, the corresponding fractional release for SO₂ and NO_x are substantially larger, 5% for SO₂ and as much as 10–15% for NO_x (refs 9, 28). Thus there is good reason for considering policy measures aimed at reducing the relatively large pollutant emissions from ships, as initiated by the International Maritime Organization mainly for coastal regions⁹. However, reduced SO₂ emissions can decrease the Earth's albedo, while reduced NO emissions from ships lead to lower concentrations of OH[•] over the open ocean, thereby enhancing the concentrations of the greenhouse-gas CH₄ and reducing the production rates of new aerosol particles. All three of these factors exert a warming effect on the Earth's climate, presenting policy makers with a striking dilemma. □

Methods

We use the Model of Atmospheric Transport and Chemistry, Max-Planck-Institute for Chemistry version¹³ (MATCH-MPIC). The meteorology component of MATCH (ref. 14 and references therein) simulates advective transport, convection, vertical diffusion, cloud fractions and microphysics. The simulations are done 'semi-offline', that is, gridded values for basic meteorological parameters (pressure, temperature, horizontal winds, surface heat fluxes, and surface stresses) are read in from the NCEP/NCAR reanalysis project data²⁹; the remaining meteorological properties (such as water vapour and cloud transport) are computed online.

The O₃–HO_x–NO_x–CH₄–CO tropospheric photochemistry version of MATCH-MPIC¹³ simulates the following: surface sources of NO_x, CO and CH₄ (for example, industry, biomass burning); free tropospheric sources of NO_x (lightning, aircraft); the stratosphere-to-troposphere fluxes of O₃ and NO_y; photochemical and photolytic processes; surface deposition loss; heterogeneous loss of N₂O₅ on aerosols; and cloud scavenging by precipitation and settling processes. The runs start in December 1992, and are continued to the end of December 1993, allowing the first December as a 'spin-up' period to reduce dependence on the initial conditions (which are from ref. 13). The model version used here is an updated version of MATCH-MPIC 2.0 (described in ref. 13); the only notable change for the present study is that the EDGAR fossil-fuel NO_x emissions⁴ (with ship emissions scaled up to 3 Tg N yr⁻¹) have replaced the previously used distribution³.

Received 20 May; accepted 15 September 1999.

- Kasibhatla, P. S., Levy, H. & Moxim, W. J. Global NO_x, HNO₃, PAN, and NO_y distributions from fossil fuel combustion emissions: A model study. *J. Geophys. Res.* **98**, 7165–7180 (1993).
- Crutzen, P. J. & Zimmermann, P. H. The changing photochemistry of the troposphere. *Tellus* **43A**–B, 136–151 (1991).
- Benkovitz, C. M. *et al.* Global gridded inventories of anthropogenic emissions of sulfur and nitrogen. *J. Geophys. Res.* **101**, 29239–29253 (1996).
- Olivier, J. G. *et al.* Description of EDGAR Version 2.0. A Set of Global Emission Inventories of Greenhouse Gases and Ozone-depleting Substances for all Anthropogenic and Most Natural Sources on a Per Country Basis and on 1° × 1° Grid (Rep. no. 771060 002 RIVM (National Institute of Public Health and the Environment, Bilthoven, 1996).
- Crutzen, P. J. & Andreae, M. O. Biomass burning in the tropics: Impact on atmospheric chemistry and biogeochemical cycles. *Science* **250**, 1669–1678 (1990).
- Yienger, J. J. & Levy, H. Empirical model of global soil-biogenic NO_x emissions. *J. Geophys. Res.* **100**, 11447–11464 (1995).
- Baughcum, S. L., Henderson, S. C., Hertel, P. S., Maggiora, D. R. & Oncina, C. A. *Stratospheric Emissions Effects Database Development* (CR-4592, NASA, Washington DC, 1994).

- Lawrence, M. G., Chameides, W. L., Kasibhatla, P. S., Levy, H. & Moxim, W. in *Handbook of Atmospheric Electrodynamics* (ed. Volland, H.) 189–202 (CRC, Boca Raton, 1995).
- Corbett, J. J., Fischbeck, P. S. & Pandis, S. N. Global nitrogen and sulfur inventories for oceangoing ships. *J. Geophys. Res.* **104**, 3457–3470 (1999).
- Carlton, J. S. *et al.* *Marine Exhaust Emissions Research Programme* (Lloyd's Register Engineering Services, London, 1995).
- Dataset of Ships 100 GRT or Greater* (Lloyd's Maritime Information Services, Stamford, Connecticut, 1996).
- Malony, M. J. *World Energy Database* (Energy Information Administration (EIA), Washington DC, 1996).
- Lawrence, M. G., Crutzen, P. J., Rasch, P. J., Eaton, B. E. & Mahowald, N. M. A model for studies of tropospheric photochemistry: Description, global distributions, and evaluation. *J. Geophys. Res.* (in the press).
- Rasch, P. J., Mahowald, N. M. & Eaton, B. E. Representations of transport, convection, and the hydrologic cycle in chemical transport models: Implications for the modeling of short lived and soluble species. *J. Geophys. Res.* **102**, 28127–28138 (1997).
- The Times Atlas of the World Comprehensive Edition* 9th edn (Times Books, London, 1992).
- International Maritime Organization (IMO) *IMO Sub-Committee on Bulk Materials, 22nd Session, Agenda Item 7* (Ref. code IMO, BCH 22/INF.10, International Maritime Organization, London, 1992).
- Houweling, S., Dentener, F. & Lelieveld, J. The impact of nonmethane hydrocarbon compounds on tropospheric photochemistry. *J. Geophys. Res.* **103**, 10673–10696 (1998).
- Sander, R. & Crutzen, P. J. Model study indicating halogen activation and ozone destruction in polluted air masses transported to the sea. *J. Geophys. Res.* **101**, 9121–9138 (1996).
- Liu, S. C. *et al.* Ozone production in the rural troposphere and the implications for regional and global ozone distributions. *J. Geophys. Res.* **92**, 4191–4207 (1987).
- Gerbig, C. *et al.* Fast response resonance fluorescence CO measurements aboard the C-130: Instrument characterization and measurements made during North Atlantic Regional Experiment 1993. *J. Geophys. Res.* **101**, 29229–29238 (1996).
- Emmons, L. K. *et al.* Climatologies of NO_x and NO_y: A comparison of data and models. *Atmos. Environ.* **31**, 1851–1904 (1997).
- Prinn, R. G. *et al.* Atmospheric trends and lifetime of CH₃CCl₃ and global OH concentrations. *Science* **269**, 187–192 (1995).
- Krol, M., van Leeuwen, P. J. & Lelieveld, J. Global OH trend inferred from methylchloroform measurements. *J. Geophys. Res.* **103**, 10697–10711 (1998).
- DeMore, W. B. *et al.* *Chemical Kinetics and Photochemical Data for Use in Stratospheric Modeling* (NASA JPL (Jet Propulsion Laboratory), Pasadena, California, 1997).
- Raes, F. Entrainment of free tropospheric aerosols as a regulating mechanism for cloud condensation nuclei in the remote marine boundary layer. *J. Geophys. Res.* **100**, 2893–2903 (1995).
- Capaldo, K., Corbett, J. J., Kasibhatla, P., Fischbeck, P. & Pandis, S. N. Effects of ship emissions on sulphur cycling and radiative climate forcing over the ocean. *Nature* **400**, 743–746 (1999).
- Hobbs, P. V. *et al.* Emissions from ships with respect to their effects on clouds. *J. Atmos. Sci.* (in the press).
- Houghton, J. T. *et al.* (eds) *Climate Change 1994* (Cambridge Univ. Press, 1995).
- Kalnay, E. *et al.* The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Met. Soc.* **77**, 437–471 (1996).

Acknowledgements

We thank P. Rasch and B. Eaton for MATCH support, and J. Olivier for information on the EDGAR ship emissions dataset. This work was supported by the EU SINDICATE project.

Correspondence and requests for materials should be addressed to M.G.L. (e-mail: lawrence@mpch-mainz.mpg.de).

A polar vortex in the Earth's core

Peter Olson & Jonathan Aurnou

Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland 21218, USA

Numerical dynamo models have been successful in explaining the origin of the Earth's magnetic field and its secular variation by convection in the electrically conducting fluid outer core^{1–7}. An important component of the convection in the numerical dynamos are polar vortices beneath the core–mantle boundary in each hemisphere. These polar vortices in the outer core have been proposed as sources for both the anomalous rotation of the inner core and the toroidal part of the geomagnetic field^{2,8}. Here we use the observed structure of the Earth's magnetic field and its variation since 1870 to infer the existence of an anticyclonic polar vortex with a polar upwelling in the northern hemisphere of the core, consistent with the polar vortices found in numerical dynamos.

The Earth's solid inner core has an important effect on the pattern

of convection in the liquid outer core and on the geomagnetic field, according to dynamo theory. Because the convection is expected to be quasi-geostrophic and nearly two-dimensional, the outer core is divided into two regions, separated by the tangent cylinder, an imaginary cylinder tangent to the inner-core equator and parallel to the Earth's rotation axis. Inside the tangent cylinder, convection-driven dynamos show intense polar vortices associated with upwellings or downwellings in both the northern and southern hemispheres. Convective upwellings that generate anticyclonic polar-vortex motion beneath the core–mantle boundary (CMB) seem to be preferred^{1,2,5,6}, but downwellings with cyclonic polar-vortex motion beneath the CMB are seen in some numerical dynamos³.

The geomagnetic field and its secular variation can be used to map the flow in the outer core beneath the CMB. Because of high electrical conductivity in the core, advection of the magnetic field by the flow dominates over diffusion and the magnetic flux lines tend to move with the fluid. This 'frozen flux' property has been used previously to construct global maps of the large-scale core flow^{9–14}. Here we use the same basic technique to delineate a more localized flow structure: a polar vortex.

Figures 1a and c show maps of the radial component of the geomagnetic field on the CMB at epochs 1870 and 1990, respectively¹⁵. The effect of an anticyclonic vortex can be seen in the change in field structure with time in the north polar region. Anticyclonic motion inside the tangent cylinder is evident from the clockwise motion of the patch of reversed magnetic flux during the

120-year interval. The boundary of this patch (the curve along which the radial flux is zero) represents a material curve in a hypothetical perfectly conducting core¹⁶. Although the core is not a perfect conductor, its conductivity is high enough that the westward rotation of the patch about the north pole implies anticyclonic flow just below the CMB. A simple cross-correlation of the fields at the two epochs indicates anticyclonic motion inside the tangent cylinder, with an average angular speed of about 0.25° per year.

Evidence for a generally southward meridional flow inside the tangent cylinder comes from the change in energy density of the polar magnetic field with time. Although the geomagnetic field is largely an axial dipole on the Earth's surface, the radial component of the field on the CMB is maximum not at the pole, as for an axial dipole, but instead at two flux bundles located just outside the tangent-cylinder region¹⁷, as shown in Fig. 1a and c. There is a minimum in intensity of the axisymmetric part of the field at the pole, and a maximum just outside the tangent cylinder. During this time period, the azimuthally averaged magnetic-energy density close to the pole tends toward zero, while the energy density near the tangent cylinder increases, as shown in Fig. 1b. This behaviour indicates that magnetic flux is being expelled from the polar region and concentrated at lower latitudes. The flux expulsion suggests a polar upwelling and southward meridional flow, which transports the field lines away from the pole and towards the tangent cylinder.

The axisymmetric part of the vortex is determined by inverting the radial component of the magnetic induction equation in the

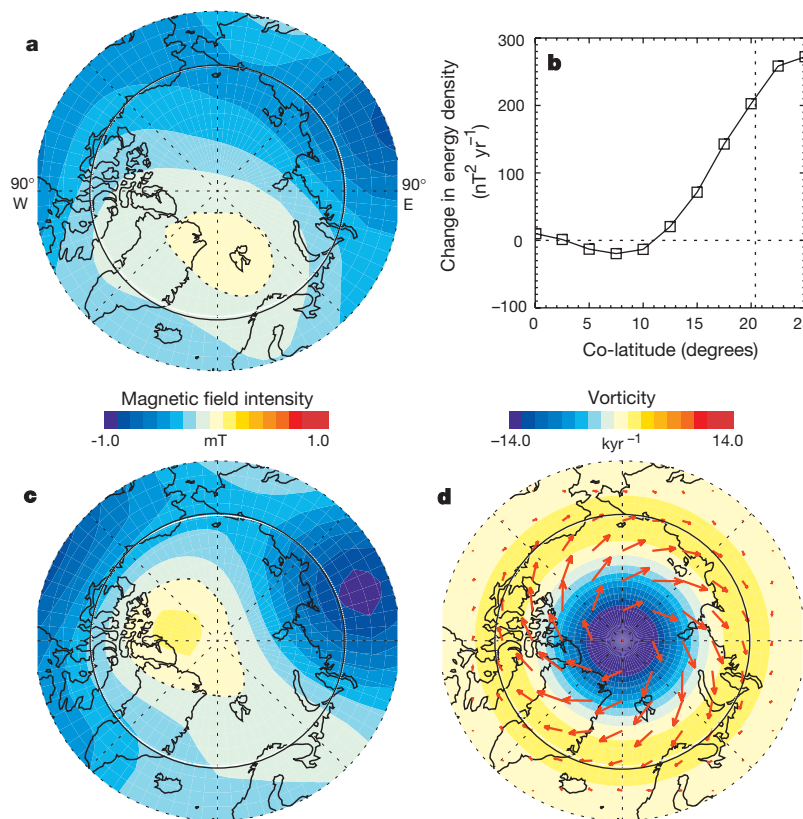


Figure 1 Secular variation of the geomagnetic field observed over the north-polar region of the core–mantle boundary and structure of the inferred axisymmetric polar vortex in the outer core. The continental outlines and the inner-core tangent-cylinder radius are shown for reference in **a**, **c** and **d**. **a**, **c**, Contours of the radial component of the geomagnetic field intensity on the core–mantle boundary in millitesla at epochs 1870 (**a**) and 1990 (**c**) from the geomagnetic field model of Bloxham and Jackson¹⁵. The westward motion of the null flux curve (dashed contours) shows the direction of fluid motion in the anticyclonic

vortex. **b**, Change in the energy density of the radial component of the magnetic field on the core–mantle boundary, in nanotesla squared per year, during 1870–1990. The energy decrease near the pole and the corresponding energy increase near the inner-core radius (dashed line near co-latitude 20.4°) imply magnetic-flux expulsion by a north-polar upwelling and generally southward meridional flow. **d**, Inferred axisymmetric structure of the north-polar vortex. Arrows indicate tangential velocities (maximum 0.13 mm s^{-1}) and colour contours indicate vorticity.

frozen-flux (high magnetic Reynolds number) limit:

$$\frac{\partial B_r}{\partial t} + \nabla_H \cdot (B_r \mathbf{U}_H) = 0 \quad (1)$$

For B_r , the radial component of the geomagnetic field, we use the time-average field on the CMB from 1870–1990 (ref. 15) and we use the difference between the fields at 1990 and 1870 to compute the secular variation $\partial B_r / \partial t$. We then solve equation (1) for $\mathbf{U}_H$, the velocity field in the outer core, tangent to the CMB (the subscript H denotes tangential coordinates.) We restrict our attention to the axisymmetric part of the flow, assumed to be steady over the time interval of the calculation. This allows us to solve for the two components of the tangential velocity without making additional assumptions. The surface integral of the axisymmetric part of equation (1) yields the meridional velocity U_θ directly. The non-axisymmetric part of equation (1) is then inverted using the least-squares method for the azimuthal velocity U_ϕ (here θ denotes co-latitude and ϕ denotes east longitude).

Figures 1d, 2 and 3 show the axisymmetric flow in the northern hemisphere found with this procedure. The azimuthal flow is anticyclonic everywhere and is concentrated inside the tangent cylinder. Its maximum angular speed corresponds to a westward drift rate of about 0.6° per year. The 120-year average westward drift rate inside the tangent cylinder is 0.25° per year, almost four times greater than the global average value over the same time period (see Fig. 3). Figure 2c and d shows the axisymmetric distributions of meridional velocity and vorticity. The vorticity is negative within 15° of the pole, but becomes positive at lower latitudes, reaching a local maximum around the tangent cylinder. The vorticity reversal has the effect of confining the azimuthal circulation to the region inside the tangent cylinder, shielding the polar vortex from the rest of the core circulation. The axisymmetric meridional flow is southward inside the tangent cylinder, but becomes northward beyond the tangent cylinder. The arrows in Fig. 2c indicate the regions of upwelling and downwelling implied by the divergence of the meridional flow. The region of implied upwelling is confined to the vortex core, within about 10° of the pole. Elsewhere the

meridional flow is convergent and the implied radial motion is downwelling.

The profiles of the azimuthal velocity, meridional velocity, vorticity and angular velocity shown in Fig. 2 are similar in shape to baroclinic, convection-driven polar vortices in the numerical dynamo calculations^{2,5,6}. These profiles are also qualitatively similar to isolated vortices generated in rotating fluids¹⁸ in the laboratory and, apart from the difference in sign, are broadly similar to the structure of the polar vortices in the stratosphere¹⁹ and to the convective chimneys in the deep ocean²⁰. The intensity of the polar vortices in the numerical dynamos tends to be larger than in Fig. 2, depending on the parameters of the calculation. But the core vortex could be stronger than our frozen-flux method indicates. When we apply our technique to the dynamo models, we recover the shapes of the polar vortices but underestimate their intensity. This is because a significant part of the flow in the dynamo models is parallel to the magnetic field contours and does not contribute to the secular variation of the field²¹. The same situation could occur in the Earth's core. We could add an arbitrary amount of motion parallel to the contour lines in Fig. 1a and c, which would not be detected using the frozen-flux technique.

Because the vortex occupies a small portion of the CMB, it is near the limit of resolution of current geomagnetic field models and near the limit of validity of the frozen-flux assumption. Accordingly, it is important to determine how robust the results are with respect to variations in model parameters. Figure 3 shows the sensitivity of the average westward drift in the vortex to variations in the epoch and window length since 1870, the year which marks the advent of polar data in the geomagnetic field model¹⁵. Figure 3 indicates that the anticyclonic motion is a persistent feature in the field model since 1870, at all epochs and for all window lengths. The shortest time windows produce a cosine fluctuation in westward drift about the 120-year average, with a period of about 80 years and a minimum near epoch 1920. Longer windows systematically approach the 120-year average. We have also calculated the axisymmetric polar flow using this geomagnetic field model truncated at spherical harmonic degree 4, the twentieth-century degree-4 field model of Yokoyama and Yukutake²², and a degree-4 archeomagnetic field model (Y. Hamano, personal communication). All three lower-resolution models show westward flow in the northern tangent cylinder.

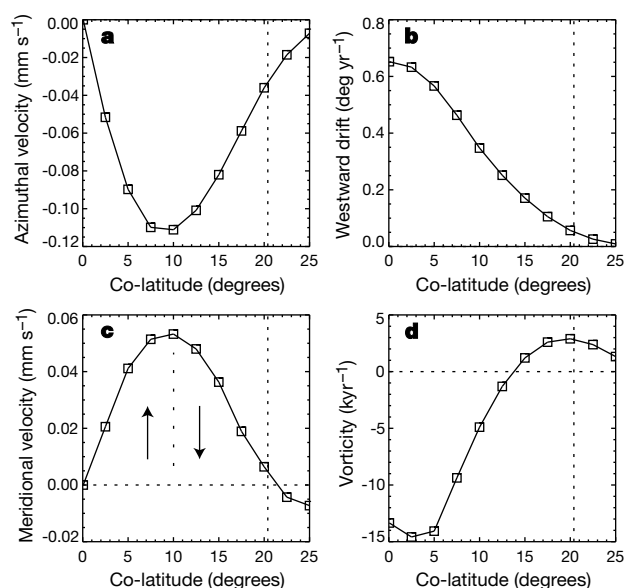


Figure 2 Profiles of the axisymmetric flow versus co-latitude in the inferred north-polar vortex 1870–1990. **a**, Azimuthal velocity (negative values are westward); **b**, angular velocity expressed as westward drift; **c**, meridional velocity (positive values are southward); **d**, vorticity (curl of the azimuthal velocity). The regions with inferred upwelling and downwelling are indicated by the arrows in **c**. Dashed lines at 20.4° indicate the co-latitude of the inner-core tangent cylinder.

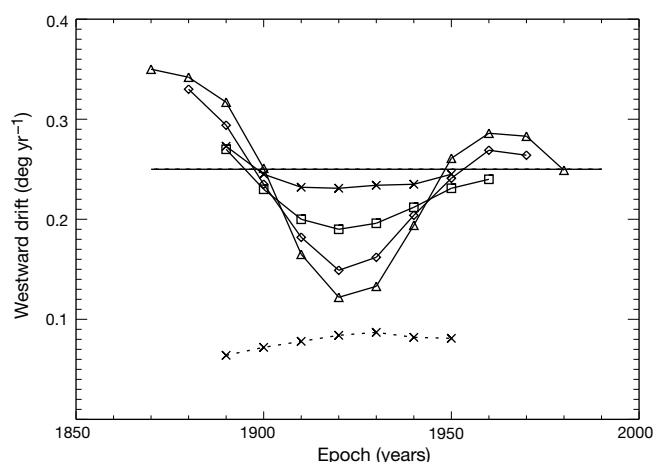


Figure 3 Persistence of the north polar vortex since 1870, determined using different time windows of geomagnetic secular variation. Solid curves show the average angular velocity inside the inner-core tangent cylinder, expressed as westward drift rate in degrees per year relative to the crust and mantle. The symbols indicate the midpoint in time (epoch) of the different time windows used: triangles, 20 years; diamonds, 40 years; squares, 60 years; crosses, 80 years. The horizontal solid line is the value for 1870–1990 from the average of Fig. 2b, obtained from the data in Fig. 1. The dotted curve shows the average drift rate over the entire core–mantle boundary.

In another test, we compare the observed and predicted geomagnetic secular variation, obtained by integrating the finite difference version of equation (1) forward in time using the velocities in Fig. 2a and c, and the 1870 field as the initial condition. A small amount of tangential diffusion was included to stabilize these calculations. The results are insensitive to diffusion for magnetic diffusivities less than $30 \text{ m}^2 \text{ s}^{-1}$ (about ten times the value estimated for the outer core²³, which supports use of the frozen-flux approximation for the axisymmetric part of the flow). The misfit between the calculated and the observed secular variation, measured using the parameter σ of Voorhies¹⁰, varies with time from ~ 0.3 to 0.5 during the 120-year interval. Overall, the axisymmetric vortex accounts for about 60% of the variance in the tangent cylinder field during this time. Nearly all of the unexplained secular variation comes from localized intensification of two features—the reversed flux patch inside the tangent cylinder and the normal-polarity flux patch beneath Siberia. The field intensification within these patches violates the frozen-flux constraint. It indicates the presence of convection on scales smaller than the field-model resolution and radial magnetic diffusion in these regions.

A frozen-flux analysis of the geomagnetic field near the south pole shows some evidence for flux expulsion there. But we do not find evidence of coherent vortex motion in the southern-hemisphere tangent cylinder. Differences in magnetic field behaviour between the polar regions might simply be an artefact of poorer data coverage near the south pole before the first decades of this century. Alternatively, it might reflect genuine differences in the flow pattern of the two regions, isolated from each other by the solid inner core.

Our results differ from most global models of the core flow. Some global models do show retrograde azimuthal motion inside the northern tangent cylinder^{9–14}, but without the large angular-velocity concentration that we find. This difference might be due to model resolution: global models use spherical harmonics and tend to fit the secular variation at low and middle latitudes, whereas our method treats only the polar region. The only exception is the Pais and Hulot global model¹¹, which focuses on the axisymmetric azimuthal part of the flow. They obtain an azimuthal flow pattern similar to Fig. 2 with a distinct polar vortex, but with larger amplitude. However, they attribute the vortex to a particular zonal spherical harmonic of the flow—degree 7—implying a similar vortex near the south pole, for which we do not find evidence.

Anticyclonic polar vorticity near the CMB has implications for the structure of convection deeper in the core. Some studies using seismic body waves have inferred anomalous prograde inner-core rotation^{24,25}, while other studies based on normal modes find no anomalous inner-core rotation relative to the crust and mantle²⁶. Either result indicates that the anticyclone weakens or possibly reverses direction somewhere between the CMB and the inner core. Previously the westward geomagnetic drift has been interpreted in terms of either cylindrically symmetric¹² or spherically symmetric²⁷ flow. Our results indicate that the convection in the north polar region has a structure more like a tropical hurricane²⁸, cylindrical in shape with circulation and vorticity changing with depth through the outer core. □

Received 30 April; accepted 28 September 1999.

1. Glatzmaier, G. A. & Roberts, P. H. A three-dimensional self-consistent computer simulation of a geomagnetic reversal. *Nature* **377**, 203–208 (1995).
2. Glatzmaier, G. A. & Roberts, P. H. Rotation and magnetism of Earth's inner core. *Science* **274**, 1887–1891 (1996).
3. Kuang, W. & Bloxham, J. An Earth-like numerical dynamo model. *Nature* **389**, 371–374 (1997).
4. Busse, F. H., Grote, E. & Tilgner, A. On convection driven dynamos in rotating spherical shells. *Stud. Geophys. Geodaet.* **42**, 1–6 (1998).
5. Olson, P., Christensen, U. & Glatzmaier, G. A. Numerical modeling of the geodynamo: Mechanisms of field generation and equilibration. *J. Geophys. Res.* **104**, 10383–10404 (1999).
6. Christensen, U., Olson, P. & Glatzmaier, G. A. Numerical modeling of the geodynamo: a systematic parameter study. *Geophys. J. Int.* **138**, 393–409 (1999).
7. Glatzmaier, G. A., Coe, R. S., Longre, L. & Roberts, P. H. The role of the Earth's mantle in controlling the frequency of geomagnetic reversals. *Nature* (in the press).

8. Aurnou, J. M., Brito, D. & Olson, P. L. Mechanics of inner core super-rotation. *Geophys. Res. Lett.* **23**, 3401–3404 (1996).
9. Voorhies, C. V. Steady flows at the top of Earth's core derived from geomagnetic field models. *J. Geophys. Res.* **91**, 12444–12466 (1986).
10. Voorhies, C. V. Time-varying flows at the top of Earth's core derived from definitive geomagnetic reference models. *J. Geophys. Res.* **100**, 10029–10039 (1995).
11. Pais, A. & Hulot, G. Core surface zonal flows, length of day decade variations and inner core superrotation. *Phys. Earth Planet. Inter.* (submitted).
12. Jault, D., Gire, C. & LeMouél, J.-L. Westward drift, core motions and exchanges of angular momentum between core and mantle. *Nature* **333**, 353–356 (1988).
13. Bloxham, J. The steady part of the secular variation of the Earth's magnetic field. *J. Geophys. Res.* **97**, 19565–19579 (1992).
14. Jackson, A. Time-dependency of tangentially geostrophic core surface motions. *Phys. Earth Planet. Inter.* **103**, 293–311 (1997).
15. Bloxham, J. & Jackson, A. Time-dependent mapping of the geomagnetic field at the core-mantle boundary. *J. Geophys. Res.* **97**, 19357–19564 (1992).
16. Backus, G. E. Kinematics of the secular variation in a perfectly conducting core. *Phil. Trans. R. Soc. Lond. A* **263**, 239–266 (1968).
17. Gubbins, D. & Bloxham, J. Morphology of the geomagnetic field and implications for the geodynamo. *Nature* **325**, 509–511 (1987).
18. Hopfinger, E. J. & Van Heijst, G. J. F. Vortices in rotating fluids. *Annu. Rev. Fluid Mech.* **25**, 241–289 (1993).
19. Schloebel, M. R. & Hartmann, D. L. The dynamics of the stratospheric polar vortex and its relation to springtime ozone depletions. *Science* **251**, 46–52 (1991).
20. Jones, H. & Marshall, J. Convection with rotation in a neutral ocean: a study of open-ocean deep convection. *J. Phys. Oceanogr.* **23**, 1009–1039 (1993).
21. Rau, S., Christensen, U., Jackson, A. & Wicht, J. Core flow inversion tested with numerical dynamo models. *Geophys. J. Int.* (submitted).
22. Yokoyama, Y. & Yukutake, T. Sixty year variation in a time series of the geomagnetic Gauss coefficients. *J. Geomagn. Geoelectr.* **43**, 563–584 (1997).
23. Stacey, F. D. *Physics of the Earth* (Brookfield, Brisbane, Australia, 1992).
24. Song, X. & Richards, P. G. Seismological evidence for the differential rotation of the Earth's inner core. *Nature* **382**, 221–224 (1996).
25. Creager, K. C. Inner core rotation rate from small-scale heterogeneity and time-varying travel times. *Science* **287**, 1284–1288 (1997).
26. Laske, G. & Masters, G. Limits on differential rotation of the inner core from an analysis of the Earth's free oscillations. *Nature* **402**, 66–69 (1999).
27. Voorhies, C. V. Inner core rotation from geomagnetic westward drift and a stationary spherical vortex in Earth's core. *Phys. Earth Planet. Inter.* **112**, 111–124 (1999).
28. Wu, C.-C. & Emanuel, K. E. Interaction of a baroclinic vortex with background shear: application to hurricane movement. *J. Atmos. Sci.* **50**, 62–76 (1993).

Acknowledgements

We thank J. Bloxham, Y. Hamano, and Y. Yokoyama for providing the geomagnetic field models and D. Waugh for pointing out similarities with other geophysical vortices. A. Pais, G. Hulot, G. Laske and G. Masters kindly shared their results in advance of publication.

Correspondence and requests for materials should be addressed to P.O. (e-mail: olson@gibbs.eps.jhu.edu).

The insidious effect of diatoms on copepod reproduction

A. Miralto*, G. Barone†, G. Romano*, S. A. Poulet‡, A. Ianora*, G. L. Russo§, I. Buttino*, G. Mazzarella§, M. Laabir*, M. Cabrini|| & M. G. Giacobbe¶

* Stazione Zoologica "A. Dohrn", Villa Comunale, 80121 Naples, Italy

† Dipartimento di Chimica Organica e Biologica, Università di Napoli "Federico II", Via Mezzocannone 16, 80134 Naples, Italy

‡ Station Biologique de Roscoff, CNRS, Place Georges Teissier, 29682 Roscoff, France

§ Istituto di Scienze dell'Alimentazione, CNR, 83100 Avellino, Italy

|| Laboratorio di Biologia Marina, Via August Piccard 54, 34040 Trieste, Italy

¶ Istituto Sperimentale Talassografico CNR, Via San Ranieri 86, 98122 Messina, Italy

The productive regions of the ocean are characterized by seasonal blooms of phytoplankton which are generally dominated by diatoms. This algal class has, therefore, traditionally been regarded as providing the bulk of the food that sustains the marine food chain to top consumers and important fisheries. However, this beneficial role has recently been questioned on the

basis of laboratory studies showing that although dominant zooplankton grazers such as copepods feed extensively on diatoms, the hatching success of eggs thus produced is seriously impaired¹. Here we present evidence from the field showing that the hatching success of wild copepods feeding on a diatom-dominated bloom is also heavily compromised, with only 12% of the eggs hatching compared with 90% in post-bloom conditions. We report on the structure of the three aldehydes isolated from diatoms that are responsible for this biological activity, and show that these compounds arrest embryonic development in copepod and sea urchin bioassays and have antiproliferative and apoptotic effects on human carcinoma cells.

In the North Adriatic, Eastern Mediterranean Sea, major diatom blooms are known to persist for several weeks in late winter². During two such blooms in 1997 and 1998, diatom-cell numbers were an order of magnitude higher than during post-bloom conditions in June (Fig. 1, lower right panel). Two species, *Skeletonema costatum* (66.3%) and *Pseudo-nitzschia delicatissima* (23.2%), dominated the bloom in 1997, whereas the following February the bloom was almost entirely composed of the diatom *P. delicatissima* (80.6%). This augmented production of diatoms triggered an increase in egg production rates of the copepod *Acartia clausi*, but average egg viability in these periods was only 12% and 24.3%, respectively, of the total number of eggs produced, compared with 90% in June (Fig. 1, upper and central right panels). Egg viability was even lower in the copepod *Calanus helgolandicus* during the bloom in 1998, when only 11.6% of the total eggs produced (9.9 ± 4.8 ; data not shown) developed to hatching. High egg mortalities have been reported for other oceanic areas³, but such losses have generally been considered to be due to the sinking of eggs to the sea floor and/or diseases, viruses and ectoparasites.

We have now isolated three antiproliferative compounds induc-

ing low hatching rates from the diatoms *Thalassiosira rotula*, *S. costatum* and *P. delicatissima*. The compounds were identified by spectroscopic techniques (high-resolution mass spectrometry; ¹H and ¹³C NMR; H–H correlation spectroscopy (COSY) and C–H COSY; ultraviolet–visual) as 2-*trans*-4-*cis*-7-*cis*-decatrinal (aldehyde A1); 2-*trans*-4-*trans*-7-*cis*-decatrinal (aldehyde A2); and 2-*trans*-4-*trans*-decadienal (aldehyde A3) (Fig. 2a). High-resolution mass spectra of aldehydes A1 and A2 gave a molecular ion at *m/e* 150.1044 and 150.1045, corresponding to a molecular formula of C₁₀H₁₄O, whereas compound A3 gave a molecular ion at 152.1197, corresponding to a molecular formula of C₁₀H₁₆O (for details, see Supplementary Information). Although all three aldehydes are known from other sources⁴, this is the first time that they have been isolated from marine diatoms to our knowledge. Aldehydes A1 and A2 have already been isolated from the freshwater diatom *Melosira varians*⁵. Diatoms synthesize unsaturated aldehydes from fatty acids through different enzymatic pathways⁵, as also reported for higher marine plants⁶.

The anti-cell-growth activity of diatom aldehydes was tested on three animal models. All three aldehydes reduced egg hatching rates in copepods (Fig. 2b) and inhibited cleavage of sea urchin embryos (Fig. 2c) when concentrations were $\geq 0.5 \mu\text{g ml}^{-1}$. In Caco2 cell lines derived from a human colon adenocarcinoma⁷, the concentration required to arrest cell growth was $11\text{--}17 \mu\text{g ml}^{-1}$ (Fig. 2d). Similar concentrations of the fatty acid *cis*-5,8,11,14,17-eicosapentaenoic acid (EPA), an inactive related compound used as a negative control because of its presence in diatoms⁵, had no effect on cell division. Figure 2e–g shows that aldehydes activated apoptotic mechanisms and that 33% of Caco2 cells were positive for nuclear staining. Aldehydes have been shown to arrest cell proliferation and induce apoptosis in cultured cell lines⁸.

These aldehydes are the probable agents of reproductive failure in

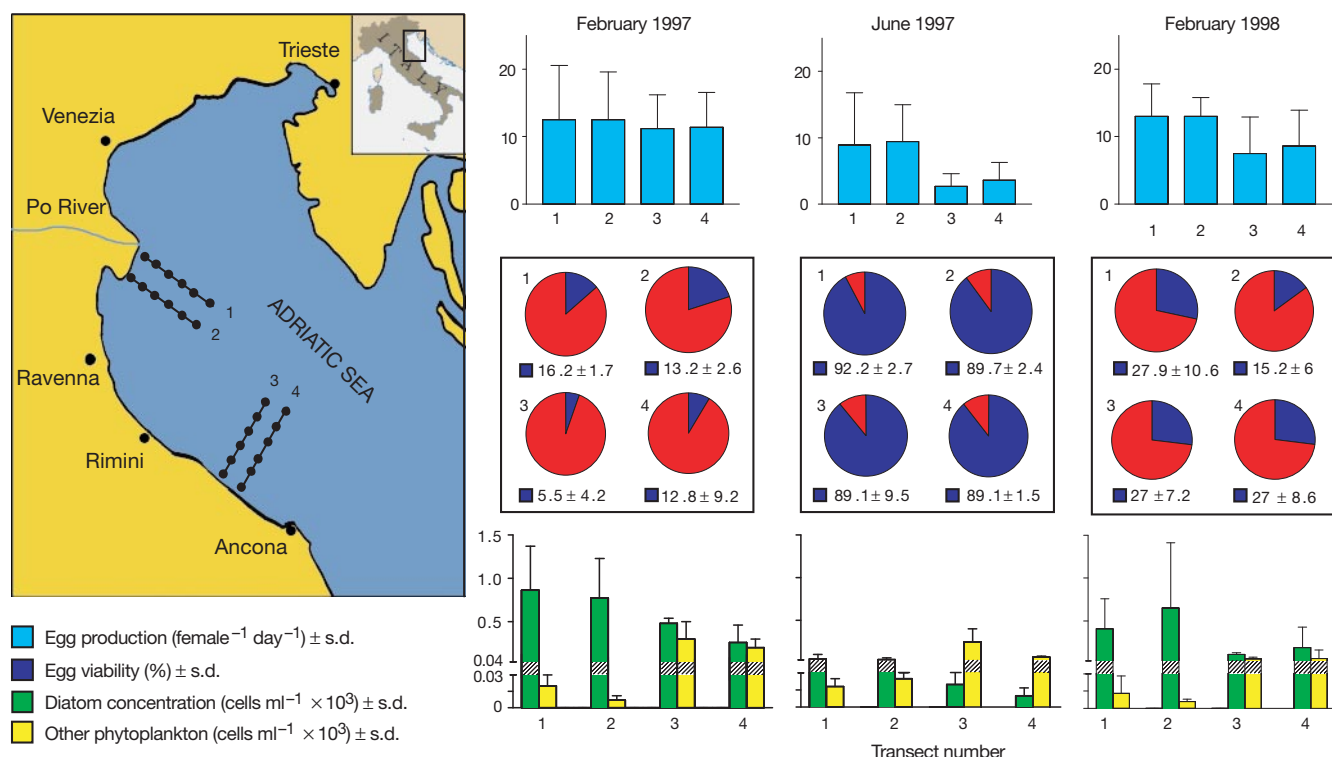


Figure 1 *In situ* relationship between diatom densities and copepod reproductive success during three cruises along four transects in the North Adriatic Sea during 1997–1998. Although the diatom bloom (green histograms) in February of both years promoted higher copepod fecundity (turquoise histograms), corresponding values for hatching success (blue portion of pie diagrams) were minimal in these periods compared with post-bloom

conditions in June when the microbial food web was established, as found for other oceanic areas¹⁶. Statistical analysis of the data showed that there was a highly significant correlation between the mean values of percentage egg mortality and diatom concentrations ($y = 1.49 + 50 \log x$, $R = 0.91$, $n = 12$; $P = 0.001$).

copepods when diatoms are the major food source. As copepods have been successfully reared from larval to adult stage on diatom food⁹, the detrimental effect on embryos but not somatic cells indicates that there may be greater vulnerability of the embryos and/or a physiological mechanism that inadvertently concentrates the aldehydes in the gonads. Some plant secondary metabolites can be much more detrimental to the embryo than the mother (for example, nicotine in *Homo sapiens*), although the role of such insidious, compared with straightforward, toxins in the evolutionary arms race between plants and herbivores is not clear. In terrestrial environments, there are many reports of substances produced by plants that interfere with the reproductive capacity of grazing animals and which act as a form of population control¹⁰. For example, ingestion of certain phytoecdysones inhibits larval growth and development in several species of insects¹¹, as well as interfering with the breeding of domestic sheep and some wild bird populations¹². Such anti-herbivory agents are thus probably common in both aquatic and terrestrial habitats even though most of the attention on plant–animal interactions until now has focused on feeding deterrents and poisoning compounds¹⁰.

Copepods can reproduce quite successfully on a diet of dinoflagellates containing potent neurotoxins (saxitoxin)¹³; indeed, wild copepods accumulate concentrations sufficient to kill their predators¹⁴. Hence, even though blooms of toxic algae occur

sporadically, copepods have evolved physiological defence systems in response to their universal toxins. In contrast, diatom blooms are a regular feature of marine seasonal cycles and copepods have silica-strengthened mandibles to crack their frustules, so it appears strange that they have not concomitantly developed a mechanism to protect their embryos. Possibly, the ability to cope with straightforward toxins evolves faster than the response to insidious ones.

Diatom blooms are widely regarded as the base of the food chain involving copepods and leading to fish. Closer examination of plankton seasonal cycles reveals, however, that the bulk of annual copepod recruitment occurs in the post-bloom phase^{15,16}. The mixed protistan community ('microbial loop') that succeeds the bloom is proposed to be a better food source for larval stages¹⁷. In this case, the effect of diatom secondary metabolites on copepod embryos would not be as detrimental as that of a universal toxin; hence evolutionary pressure to develop antidotes against the 'birth control effect' of the diatoms would not build up to the same extent. □

Methods

Purification and identification of aldehydes

Procedures for culturing axenic *T. rotula* cells are given elsewhere¹⁸; the same procedures were also used to culture non-axenic *S. costatum* and *P. delicatissima*. Extracts were prepared by centrifuging cells in exponential growth stage for 10 min at 1500g at 4 °C. About 18×10^9 pelleted cells (~100 g total wet weight) were sonicated in an ice bath with 100 ml of bi-distilled water and centrifuged at 75,000g for 2 h at 4 °C. The supernatant was extracted four times with an equal volume of dichloromethane. The pooled active organic residue was chromatographed in two steps using silica gel 60 F₂₅₄ TLC plates (0.5 mm, Merck) with hexane/ethyl acetate (9/1) (two runs) and a C-8 bonded-phase silica gel F₂₅₄ plate (Merck 0.25 mm) with hexane/ethyl acetate (95/5) as eluents (two runs).

Copepod, sea urchin and Caco2 cell bioassays

Fertilized eggs were freely spawned by female *Temora stylifera* copepods incubated in 5-ml tissue-culture wells containing aldehydes A1–A3 or the fatty acid EPA at final concentrations ranging from 0.5–2.0 $\mu\text{g ml}^{-1}$; the percentage of hatched eggs was determined 48 h later when control untreated eggs had hatched. The same aldehyde and EPA concentrations were also tested on sea urchin *Paracentrotus lividus* embryos; the percentage of blocked embryos was calculated when control embryos were at the 8-blastomere stage. Sea urchin fertilization bioassay procedures are described elsewhere¹⁹. In typical Caco2 assays, 300,000 cells ml^{-1} were cultured in multi-well plates as described elsewhere⁷. After 48 h, cell proliferation was determined using an MTT (thiazolyl blue) assay and confirmed by cell counting in trypan blue stain. Apoptosis was verified using Tdt-mediated dUTP nick end labelling (TUNEL) (Boehringer). Forty-eight hours after addition of 5 $\mu\text{g ml}^{-1}$ of A1, floating cells were washed in phosphate buffer saline and counted using the Zeiss 410 confocal laser scanning microscope (488Å) at 40× magnification.

Field samples

At each station, female *A. clausi* and *C. helgolandicus* copepods were incubated individually in Falcon tissue culture flasks containing 80 ml seawater with natural phytoplankton. Eggs were counted after 24 h, whereas the percentage of hatched eggs was determined after an additional 48 h. Phytoplankton samples were taken at alternate stations (4 or 5 depths between 0 and 63 m) with Niskin bottles, fixed (in neutralized formaldehyde) and analysed in Utermohl settling chambers; reported cell densities are integrated for the entire water column.

Received 12 May; accepted 20 August 1999.

1. Ban, S. *et al.* The paradox of diatom–copepod interactions. *Mar. Ecol. Prog. Ser.* **157**, 287–293 (1997).
2. Vollenweider, R. A., Rinaldi, A. & Montanari, G. in *Marine Coastal Eutrophication* (eds Vollenweider, R. A., Marchetti, R. & Viviani, D.) 63–106 (Elsevier, Amsterdam, 1992).
3. Peterson, W. T. & Kimmmerer, W. J. Processes controlling recruitment of the marine calanoid copepod *Temora longicornis* in Long Island Sound: egg production, egg mortality, and cohort survival rates. *Limnol. Oceanogr.* **39**, 1594–1605 (1994).
4. Harkes, P. D. & Begemann, W. J. Identification of some previously unknown aldehydes in cooked chicken. *J. Am. Chem. Soc.* **51**, 356–359 (1974).
5. Wendel, T. & Jüttner, F. Lipoxygenase-mediated formation of hydrocarbons and unsaturated aldehydes in freshwater diatoms. *Phytochemistry* **41**, 1445–1449 (1996).
6. Gerwick, W. H. & Bernart, M. W. in *Marine Biotechnology*, Vol. 1. *Pharmaceutical and Bioactive Natural Products*. (eds Attaway, D. H. & Zaborsky, O. R.) 101–152 (Plenum, New York, 1993).
7. Fogh, J., Wright, W. C. & Loveless, J. D. Absence of HeLa cell contamination in cell lines derived from human tumors. *J. Natl Cancer Inst.* **58**, 209–214 (1977).
8. Zimmerman, B. T., Crawford, G. D., Dahl, R., Simon, F. T. & Mapoles, J. E. Mechanisms of acetaldehyde-mediated growth inhibition: delayed cell cycle progression and induction of apoptosis. *Alcoholism Clin. Exp. Res.* **19**, 434–440 (1995).
9. Koski, M., Klein Breteler, W. & Schogt, N. Effect of food quality on rate of growth and development of the pelagic copepod *Pseudocalanus elongatus* (Copepoda, Calanoida). *Mar. Ecol. Prog. Ser.* **170**, 169–187 (1998).

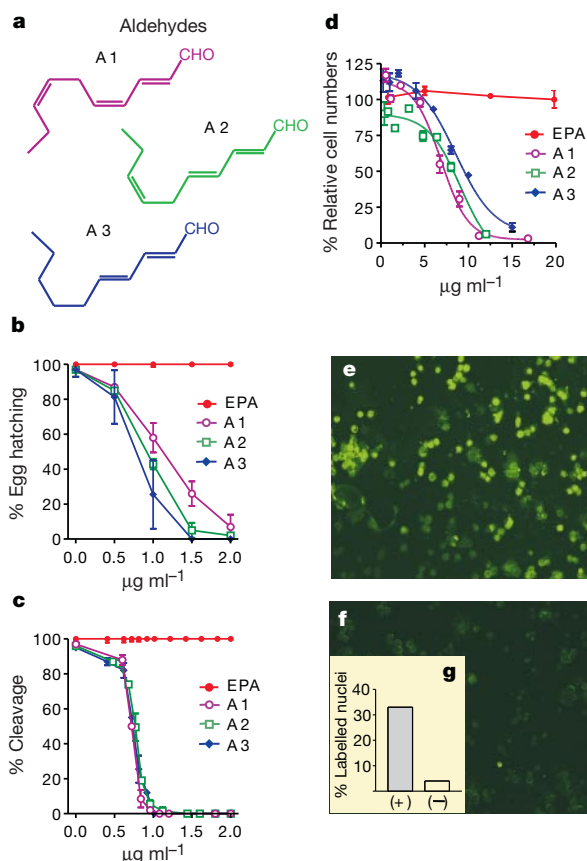


Figure 2 Effect of aldehydes on cell division. Aldehydes A1 (pink), A2 (green) and A3 (blue) (a) markedly reduced the percentage hatching success of newly spawned copepod eggs (b), cleavage of sea urchin embryos (c) and growth of Caco2 cells (d). The related compound EPA (red), used as a negative control, had no effect on cell division. Cells incubated with 5 $\mu\text{g ml}^{-1}$ of A1 (e) underwent apoptosis after 48 h of treatment compared with control untreated cultures (f). Thirty-three per cent of the nuclei were positive for nuclear staining with TUNEL (g). All results are expressed as means $\pm$ s.d. for triplicate determinations.

10. Harborne, J. B. in *Herbivores: Their Interactions with Secondary Plant Metabolites* (eds Rosenthal, G. A. & Berenbaum, M. R.) 389–429 (Academic, San Diego, 1991).
11. Robbins, W. E. *et al.* Ecdysones and analogs: Effects on development and reproduction of insect. *Science* **161**, 1158–1160 (1968).
12. Leopold, A. S. *et al.* Phytoestrogens: Adverse effects on reproduction in California quail. *Science* **191**, 98–100 (1976).
13. Dutz, J. Repression of fecundity in the meritic copepod *Acartia clausi* exposed to the toxic dinoflagellate *Alexandrium lusitanicum*: relationship between feeding and egg production. *Mar. Ecol. Prog. Ser.* **175**, 97–107 (1998).
14. White, A. W. Marine zooplankton can accumulate and retain dinoflagellate toxins and cause fish kills. *Limnol. Oceanogr.* **26**, 103–109 (1981).
15. Kleppel, G. S., Holliday, D. V. & Pieper, R. E. Trophic interactions between copepods and microplankton: a question about the role of diatoms. *Limnol. Oceanogr.* **36**, 172–178 (1991).
16. Runge, J. A. & de Lafontaine, Y. Characterization of the pelagic ecosystem in surface waters of the northern Gulf of St. Lawrence in early summer: the larval redfish-*Calanus*-microplankton interaction. *Fish. Oceanogr.* **5**, 21–37 (1996).
17. Smetacek, V. S. Role of sinking in diatom life-history cycles: ecological, evolutionary and geological significance. *Mar. Biol.* **84**, 239–251 (1985).
18. Ianora, A., Poulet, S. A., Miralto, A. & Grotto, R. The diatom *Thalassiosira rotula* affects reproductive success in the copepod *Acartia clausi*. *Mar. Biol.* **125**, 279–286 (1996).
19. Buttino, L., Miralto, A., Ianora, A., Romano, G. & Poulet, S. A. Water soluble extracts of the diatom *Thalassiosira rotula* induce aberrations in embryonic tubulin organisation of sea urchin *Paracentrotus lividus*. *Mar. Biol.* **134**, 147–154 (1999).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank E. Trivellone and T. Tancredi of the Core Facility, CNR Naples, for NMR assistance, and M. Di Pinto, F. Esposito and V. Mirra for technical help. This work was funded by the Stazione Zoologica and Roscoff Marine Station and an Italian MURST programme PRISMA II, Biogeochemical Cycles.

Correspondence and request for materials should be addressed to A.M. (e-mail: miralto@alpha.szn.it).

Signal but not noise changes with perceptual learning

J. Gold, P. J. Bennett & A. B. Sekuler

Department of Psychology, University of Toronto, 100 St George Street, Toronto, Ontario M5S 3G3, Canada

Perceptual discrimination improves with practice. This 'perceptual learning' is often specific to the stimuli presented during training^{1–5}, indicating that practice may alter the response characteristics of cortical sensory neurons^{6,7}. Although much is known about how learning modifies cortical circuits⁸, it remains unclear how these changes relate to behaviour. Different theories assume that practice improves discrimination by enhancing the signal^{1,9,10}, diminishing internal noise^{11,12} or both¹³. Here, to distinguish among these alternatives, we fashioned sets of faces and textures whose signal strength could be varied, and we trained observers to identify these patterns embedded in noise. Performance increased by up to 400% across several sessions over several days. Comparisons of human performance to that of an ideal discriminator showed that learning increased the efficiency with which observers encoded task-relevant information. Observer response consistency, measured by a double-pass technique in which identical stimuli are shown twice in each experimental session^{14,15}, did not change during training, showing that learning had no effect on internal noise. These results indicate that perceptual learning may enhance signal strength, and provide important constraints for theories of learning.

In our experiments, observers identified which of 10 possible visual signals was presented on each trial. To make identification difficult, two-dimensional gaussian noise was added to the signals

on the visual display. An ideal strategy for this task¹⁶ is to cross-correlate the stimulus (the signal-plus-noise) with 10 different templates which match the spatial and temporal characteristics of the 10 possible signals, and then select the template that yields the largest response. The performance of such an ideal observer is limited only by the similarity of the items and by the amount of noise added to the signal.

Of course, additional factors constrain human performance. For example, the optics of the eye degrade the stimulus, and neural mechanisms may further attenuate the stimulus and/or introduce noise. In addition, human observers may encode stimuli with filters that are not matched precisely to the spatial and/or temporal characteristics of the signal, or use suboptimal decision rules to select a response. These additional constraints on human performance have been modelled successfully as a two-component system containing a constant, internal noise added to the external stimulus, followed by a noiseless, contrast-invariant calculation^{17–19}. The calculation can be thought of as the application of a linear filter, or template, that reduces the input to a single number used to make a decision about the stimulus. To the extent that the template does not match the signal, the efficiency of the calculation is reduced relative to ideal efficiency¹⁷. In this model, human thresholds follow the form

$$E = k(N_e + N_i) \quad (1)$$

where E is threshold contrast energy (the squared root mean square (r.m.s.) contrast multiplied by stimulus area), N_e is the spectral density of the external noise, and k and N_i are free parameters¹⁸. N_i is an index of the total internal noise which affects performance, whereas k is inversely proportional to calculation efficiency (see Methods for a definition of calculation efficiency). Equation (1) suggests that human observers differ from ideal observers in two ways. First, unlike ideal observers, human observers have internal noise. Second, unlike ideal observers, human observers use sub-

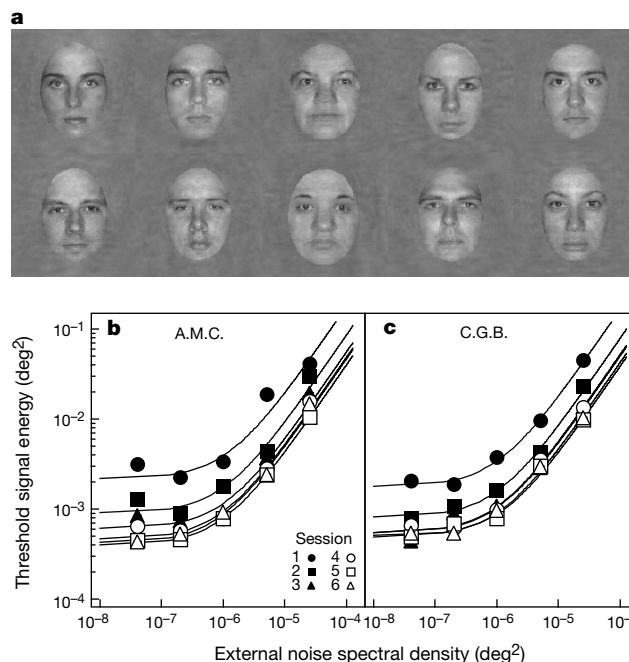


Figure 1 Faces and their noise masking functions. **a**, Stimuli used in the face identification experiment (see Methods). **b, c**, Contrast energy thresholds plotted as a function of external noise spectral density (power per unit bandwidth) for two observers in the face identification task. Each data point corresponds to a single threshold. Error bars correspond to ± 1 s.e. of the threshold estimate. Often, the error bars are smaller than the symbols. The solid lines are the weighted least-squares fits to equation (1).

optimal templates, which reduces their relative calculation efficiency. Our experiments determined whether perceptual learning lowers thresholds by decreasing internal noise or by increasing calculation efficiency.

Face identification thresholds were measured for 10 unfamiliar human faces (Fig. 1a) embedded in static, two-dimensional gaussian white noise of various spectral densities. A single, randomly selected face-plus-noise stimulus was presented on each trial. Thresholds increase as a function of external noise and decrease as a function of practice (Fig. 1b, c). The smooth curves in Fig. 1b, c are the best fitting form of equation 1 for each set of thresholds. When plotted in log-log coordinates, as in Fig. 1b, c, changes to k result in uniform vertical shifts of the curve, whereas changes to N_i result in lateral shifts of the curve's inflection point. Figure 1b, c indicates that practice reduced k but had no effect on N_i . Thus, our results indicate that perceptual learning may increase calculation efficiency without affecting internal noise. Calculation efficiency and internal noise estimated from the data in Figs 1 and 2 are shown in Fig. 3, where the circular symbols show the above result more clearly.

Previous studies have suggested that specialized mechanisms process faces^{20,21}. Therefore, our results may not be characteristic of perceptual learning in general. To test this idea, we applied the same technique to a stimulus without the spatial structure and social significance of human faces: band-pass filtered gaussian noise textures (Fig. 2a). Texture identification thresholds increased as a function of external noise and decreased as a function of practice (Fig. 2b, c). Calculation efficiency and internal noise estimated from the data in Fig. 2 are shown by the square symbols in Fig. 3. As with faces, practice increased calculation efficiency (reduced k) but had no effect on internal noise (N_i did not vary). Thus, this pattern of results is not specific to faces, and may be generally representative of how people learn to identify visual patterns.

The increase in calculation efficiency with practice is consistent with the idea that learning modifies observers' templates to extract

more of the potential signal. However, another interpretation is possible. The internal noise N_i in equation (1) does not depend on the strength of the external noise or the signal, but physiological²² and psychophysical¹⁵ evidence indicates that the variability of visual mechanisms may increase with stimulus contrast. These results imply that there are two sources of internal noise: a constant component N_c and a multiplicative component which is proportional to the input¹⁵. A decrease in multiplicative noise has the same effect as an increase in calculation efficiency (that is, a uniform downward shift in the functions shown in Figs 1b, c and 2b, c)^{15,18}. This can be seen more clearly when equation (1) is rewritten so that internal noise, N_i , is the sum of constant (N_c) and multiplicative ($m\{N_e + N_c + E\}$) components:

$$\begin{aligned} E &= k(N_e + N_c + m(N_e + N_c + E)) \\ &= (k(1 + m)/(1 - km))(N_e + N_c) \\ &= k'(N_e + N_c) \end{aligned} \quad (2)$$

where k' is a constant equal to $k(1 + m)/(1 - km)$. Note that reductions in k' can be caused either by reducing multiplicative noise (lowering the value of m) or by increasing calculation efficiency (lowering k). Thus, within the context of our model, it

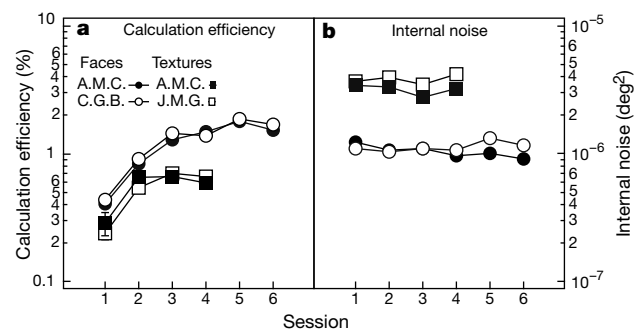


Figure 3 Calculation efficiency and internal noise estimates. **a**, Calculation efficiency; **b**, internal noise spectral density estimates from the data in Figs 1 and 2, plotted as a function of practice. Error bars correspond to ± 1 s.e. of the estimate.

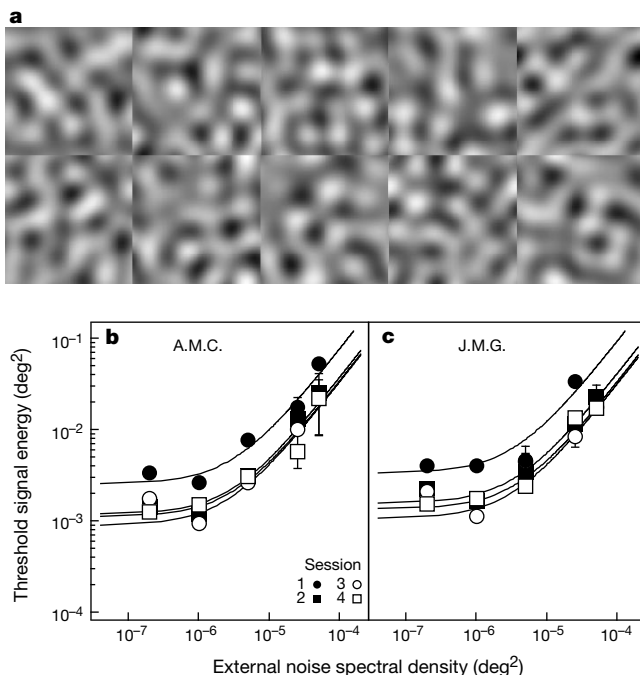


Figure 2 Band-pass texture patterns and their noise masking functions. **a**, Stimuli used in the texture identification experiment (see Methods). **b**, **c**, Contrast energy thresholds plotted as a function of external noise spectral density for two observers in the texture identification task. Observer A.M.C. (**b**) also participated in the face identification experiment (Fig. 1b). Plotting conventions are as in Fig. 1.

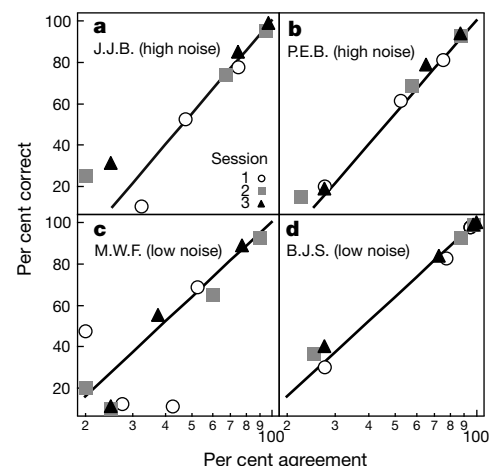


Figure 4 Response consistency in high and low external noise. High (**a**, **b**) and low (**c**, **d**) external noise consistency measures in the face identification task. Each panel plots per cent correct performance as a function of per cent agreement. Each data point corresponds to the combination of per cent correct and per cent agreement at a single stimulus level. All observers showed the same degree of improvement across test sessions as found in the first experiment. Solid lines show the predictions for an observer with an internal/external noise ratio of 1.0 (**a**, **b**) and 2.0 (**c**, **d**).

is impossible to know whether the effects of practice on identification thresholds are due to reductions in multiplicative noise or to increases in calculation efficiency.

To address this issue, we adopted a double-pass response consistency procedure which has been used to estimate internal noise for auditory¹⁴ and visual^{15,23} detection and discrimination tasks. The technique requires observers to perform a task with two passes through identical sets of stimuli (signal-plus-noise combinations). Because both the signals and noises in both passes are identical, response inconsistencies between passes must be due to observer variability. For a given level of performance (per cent correct), the degree of consistency (per cent agreement) between responses in the two repetitions of the experiment depends only on the internal/external noise ratio¹⁵. Specifically, as the internal/external noise ratio increases, response consistency decreases. Thus, measures of response consistency are an index of internal noise. Furthermore, the estimate of internal noise does not depend on whether the noise is constant or multiplicative—it simply reflects the total amount of noise present in the system. We used this technique to determine whether internal noise changes with practice in our task.

Four new observers performed the face identification task with two passes through the same signal and noise combinations within each of three sessions. The consistency measures are shown in Fig. 4. If learning reduced internal noise, consistency would increase with practice, and the data in each panel of Fig. 4 should shift systematically to the right. However, for both high and low external noise conditions, the data for all three sessions fall along a single line. Contrary to the predictions of several theories^{11–13}, internal noise was not altered by learning.

Our results provide strong constraints for models of learning, and support connectionist network models^{1,10} and classical learning theories⁹ suggesting that learning increases the strength of discriminate signals. Precisely how signal enhancement manifests itself at the neuronal level remains a mystery. Learning produces changes in individual cells^{24,25}, populations of cells within a single cortical area^{26–28} and functional connections across cortical areas²⁹. An important challenge for the future is to determine whether some or all of these physiological changes are associated with changes in behaviour. Our results suggest that the neural correlates of behaviour should exhibit improved discrimination among stimuli and constant variability during perceptual learning. □

Methods

Definition of calculation efficiency

It can be shown that identification threshold for an ideal observer is given by $E = k_{\text{ideal}} * N_e$, where E and N_e are as defined above and k_{ideal} is a constant that varies with the pattern set and is directly related to the intrinsic difficulty of the task¹⁶. A human observer's calculation efficiency is defined as k_{ideal}/k .

Stimuli

We used 10 faces (Fig. 1a) and 10 textures (Fig. 2a) as stimuli. The height/width ratio was constant across all faces (198 pixels/140 pixels), subtending $4.0 \times 2.9^\circ$ of visual angle from the viewing distance of 100 cm. The faces were centred within a 256×256 pixel ($5.25 \times 5.25^\circ$) uniform background of average luminance (29 cd m^{-2}). Additional details about the generation of the face stimuli are described elsewhere³⁰. The textures were gaussian noise fields (256×256 pixels; $5.25 \times 5.25^\circ$) filtered by a 2–4 c per image rectangular frequency filter.

Procedure

On each trial, a signal (for example, a face) was chosen and added to a two-dimensional static gaussian white noise field (256×256 pixels) of the appropriate contrast variance. The signal-plus-noise combination was shown for 500 ms and was followed by an unlimited time for response by clicking on a selection window containing smaller (100×100 pixels), noise-free, high-contrast versions of all 10 signals. Observers received accuracy feedback after each trial. Thresholds for a d' of 1.47 (50% correct) were determined by varying the contrast energy of the signals across trials according to the method of constant stimuli.

Design of first experiment

See Figs 1–3. Thresholds were determined in each of five levels of levels of external noise:

0.04, 0.20, 1.02, 5.11 and $25.55 \times 10^{-6} \text{ deg}^2$. A unique noise field was generated on every trial, and the level of external noise and the identity of the signal were chosen randomly. A complete session consisted of 155 trials across each of five stimulus contrast energy levels at each noise level, for a total of 775 trials per session. Each session was completed without breaks and lasted about 1 h. Only one session was completed each day. Observers in the face identification task completed six sessions within ten days. Observers in the texture identification task completed four sessions within seven days.

Design of double-pass consistency experiment

See Fig. 4. Two observers performed the face identification task with only the highest level of external noise used in the first experiment ($25.55 \times 10^{-6} \text{ deg}^2$), and two with only the lowest level of external noise ($0.04 \times 10^{-6} \text{ deg}^2$). Each of three test sessions consisted of two identical blocks of 200 trials across five stimulus contrast energy levels, for a total of 400 trials per session. Signals were chosen randomly and new noise fields were generated on each trial within the first block of a session. Observers were not aware that the first and second blocks of each session were identical.

Received 13 July; accepted 13 September 1999.

- Poggio, T., Fahlke, M. & Edelman, S. Fast perceptual learning in visual hyperacuity. *Science* **256**, 1018–1021 (1992).
- Florentini, A. & Berardi, N. Perceptual learning specific for orientation and spatial frequency. *Nature* **287**, 43–44 (1980).
- Fahlke, M. & Morgan, M. No transfer of perceptual learning between similar stimuli in the same retinal position. *Curr. Biol.* **6**, 292–297 (1996).
- Ball, K. & Sekuler, R. Direction-specific improvement in motion discrimination. *Vis. Res.* **27**, 953–965 (1987).
- Sagi, D. & Tanne, D. Perceptual learning: learning to see. *Curr. Opin. Neurobiol.* **4**, 195–199 (1994).
- Karni, A. & Sagi, D. Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proc. Natl Acad. Sci. USA* **88**, 4966–4970 (1991).
- Vogels, R. & Orban, G. A. The effect of practice on the oblique effect in line orientation judgments. *Vis. Res.* **25**, 1679–1687 (1985).
- Buonomano, D. & Merzenich, M. Cortical plasticity: From synapses to maps. *Annu. Rev. Neurosci.* **21**, 149–186 (1998).
- Mackintosh, N. *Conditioning and Associative Learning* (Oxford Univ. Press, New York, 1983).
- Weiss, Y., Edelman, S. & Fahlke, M. Models of perceptual learning in vernier hyperacuity. *Neural Comp.* **5**, 695–718 (1993).
- McLaren, I., Kaye, H. & Mackintosh, N. in *Parallel Distributed Processing: Implications for Psychology and Neurobiology* (ed. Morris, R.) 102–130 (Clarendon, Oxford, 1988).
- McLaren, I. in *Causal Mechanisms of Behavioural Development* (eds Hogan, J. & Bolhuis, J.) 377–402 (Cambridge Univ. Press, Cambridge, 1994).
- Doshier, B. A. & Lu, Z. Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proc. Natl Acad. Sci. USA* **95**, 13988–13993 (1998).
- Green, D. M. Consistency of auditory detection judgments. *Psychol. Rev.* **71**, 392–407 (1964).
- Burgess, A. E. & Colborne, B. Visual signal detection. IV. Observer inconsistency. *J. Opt. Soc. Am. A* **5**, 617–627 (1988).
- Tjan, B. S., Braje, W. L., Legge, G. E. & Kersten, D. Human efficiency for recognizing 3-D objects in luminance noise. *Vis. Res.* **35**, 3053–3069 (1995).
- Legge, G., Kersten, D. & Burgess, A. E. Contrast discrimination in noise. *J. Opt. Soc. Am. A* **4**, 391–406 (1987).
- Pelli, D. G. in *Vision: Coding and Efficiency* (ed. Blakemore, C.) 3–24 (Cambridge Univ. Press, Cambridge, Massachusetts, 1990).
- Burgess, A. E., Wagner, R. F., Jennings, R. J. & Barlow, H. B. Efficiency of human visual signal discrimination. *Science* **214**, 93–94 (1981).
- Desimone, R. Face-selective cells in the temporal cortex of monkeys. *J. Cog. Neurosci.* **3**, 1–8 (1991).
- Yin, R. K. Looking at upside-down faces. *J. Exp. Psychol.* **81**, 141–145 (1969).
- Tolhurst, D. J., Movshon, J. A. & Dean, A. F. The statistical reliability of signals in single neurons in cat and monkey cortex. *Vis. Res.* **23**, 775–785 (1983).
- van Meeteren, A. & Barlow, H. B. The statistical efficiency for detecting sinusoidal modulation of average dot density in random figures. *Vis. Res.* **21**, 765–777 (1981).
- Zohary, E., Celebrini, S., Britten, K. H. & Newsome, W. T. Neuronal plasticity that underlies improvement in perceptual performance. *Science* **263**, 1289–1292 (1994).
- Logothetis, N. K., Pauls, J. & Poggio, T. Shape representation in the inferior temporal cortex of monkeys. *Curr. Biol.* **5**, 552–563 (1995).
- Pascual-Leone, A. & Torres, F. Plasticity of the sensorimotor cortex representation of the reading finger in Braille readers. *Brain* **116**, 39–52 (1993).
- Recanzone, G. H., Schreiner, C. E. & Merzenich, M. M. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci.* **13**, 87–103 (1993).
- Schiltz, C. et al. Neuronal mechanisms of perceptual learning: changes in human brain activity with training in orientation discrimination. *Neuroimage* **9**, 46–62 (1999).
- McIntosh, A. R., Rajah, M. N. & Lobaugh, N. J. Interactions of prefrontal cortex in relation to awareness in sensory learning. *Science* **284**, 1531–1533 (1999).
- Gold, J., Bennett, P. J. & Sekuler, A. B. Identification of band-pass filtered letters and faces by human and ideal observers. *Vis. Res.* **39**, 3537–3560 (1999).

Acknowledgements

This research was funded by grants to P.J.B. and A.B.S. from the Natural Science and Engineering Research Council of Canada. We thank R. Sekuler for comments and suggestions.

Correspondence and requests for materials should be addressed to P.J.B. (e-mail: bennett@psych.utoronto.ca).

Conflict monitoring versus selection-for-action in anterior cingulate cortex

Matthew Botvinick^{*†}, Leigh E. Nystrom[‡], Kate Fissell[§],
Cameron S. Carter[†] & Jonathan D. Cohen^{†‡}

^{*} Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA

[†] Department of Psychiatry, University of Pittsburgh Medical Center, 3811 O'Hara St., Pittsburgh, Pennsylvania 15213, USA

[‡] Department of Psychology, Princeton University, Princeton, New Jersey 08544-1010, USA

[§] Department of Psychology, University of Pittsburgh, 405 Langley Hall, Pittsburgh, Pennsylvania 15260, USA

The anterior cingulate cortex (ACC), on the medial surface of the frontal lobes of the brain, is widely believed to be involved in the regulation of attention^{1,2}. Beyond this, however, its specific contribution to cognition remains uncertain. One influential theory has interpreted activation within the ACC as reflecting 'selection-for-action'³⁻⁵, a set of processes that guide the selection of environmental objects as triggers of or targets for action. We have proposed an alternative hypothesis, in which the ACC serves not to exert top-down attentional control but instead to detect and signal the occurrence of conflicts in information processing⁶⁻⁸. Here, to test this theory against the selection-for-action theory, we used functional magnetic resonance imaging to measure brain activation during performance of a task where, for a particular subset of trials, the strength of selection-for-action is inversely related to the degree of response conflict. Activity within the ACC was greater during trials featuring high levels of conflict (and weak selection-for-action) than during trials with low levels of conflict (and strong selection-for-action), providing evidence in favour of the conflict-monitoring account of ACC function.

Subjects performed a version of the flanker task⁹ in which they were asked to indicate by button-press the orientation of a briefly presented left- or right-facing arrow. The target arrow was always flanked by a set of distractor arrows, two on each side. On compatible trials, these pointed in the same direction as the target arrow (for example, <<<<<<); on incompatible trials, they faced in the opposite direction (for example, <<>><<).

The flanker task involves both conflict and selection-for-action, in particular on incompatible trials. On these trials the combined influence of the target and flankers leads to conflict in the form of competition between correct and incorrect responses, an effect that is reflected in prolonged reaction times⁹⁻¹¹. At the same time, attending to the target but not the flankers calls for selection-for-action, a mechanism that according to A. Allport, the term's originator, "can selectively designate a specified subset of the available, and potentially relevant, sensory information to have

control of a given effector system, and can selectively decouple the remainder from such control"¹².

Our first prediction, based on the conflict-monitoring theory, was that ACC activation would be greater on incompatible trials, as this is where conflict is greatest. However, to adjudicate clearly between the two theories of ACC function, we turned to a second prediction, based on a phenomenon we refer to as the Gratton effect.

Gratton *et al.*¹³ discovered that the balance between conflict and selection-for-action on any given trial in the flanker task depends on the compatibility of the preceding trial. Specifically, they showed (based on reaction time data) that the distracting effect of the flankers is weaker on trials that follow incompatible trials than on trials that follow compatible ones. In effect, the occurrence of an incompatible trial leads to a strengthening of selection-for-action, reducing the influence of the flankers during the subsequent trial.

The most interesting implications of the Gratton effect relate specifically to incompatible trials. For purposes of analysis, we divide these trials into two categories: those that follow compatible trials (here labelled cI), and those that follow incompatible ones (iI). Owing to the Gratton effect, these two groups of trials involve very different proportions of selection-for-action and conflict. Incompatible trials that follow other incompatible trials (iI trials) involve relatively strong selection-for-action and, as a result, diminished flanker-induced conflict. Conversely, incompatible trials coming after compatible ones (cI trials) involve relatively weak selection-for-action and, thus, more flanker-induced conflict.

These observations prompt the two theories of ACC function to make opposite predictions about brain activation during cI and iI trials. The conflict-monitoring theory predicts that ACC activation should be greater on cI trials, when conflict is at its highest. The opposing selection-for-action theory predicts that activation should be greater on iI trials, when there is stronger selection-for-action.

No task yet studied has been found to engage the entire ACC, leaving open the possibility of functional heterogeneity within the region. Our predictions therefore specifically relate to the portions of ACC to which a selection-for-action role has previously been ascribed. These centre on a region anterior to the plane of the anterior commissure, posterior to the genu of the corpus callosum and often extending into the cingulate sulcus.

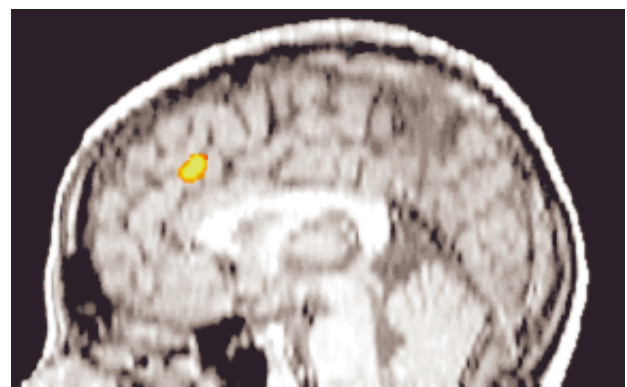


Figure 2 Location of the ACC area displaying greater activity on incompatible than compatible trials and on cI than iI trials. The area shown represents the overlap of regions identified in independent ANOVAs (see Methods), each using a significance threshold of $P = 0.01$. Peak F values for both regions fell within the area depicted (scan $\times$ trial-type interaction: $F = 17.16$, coordinates $x = -2$, $y = 28$, $z = 31$; scan $\times$ previous trial-type interaction: $F = 5.2$, coordinates $x = -2$, $y = 31$, $z = 29$). Areas outside the ACC showing an effect of current trial type with a peak F equal to or greater than that observed in the ACC included right inferior parietal lobule (BA 40) and left anterior insula. Areas showing an effect of previous trial type on incompatible trials with a peak F equal to or greater than that observed in the ACC included left dorsolateral prefrontal cortex (BA 9/46), bilateral postcentral gyrus, and bilateral inferior parietal lobe (BA 40, with extension into BA 39 and 22 on the left).

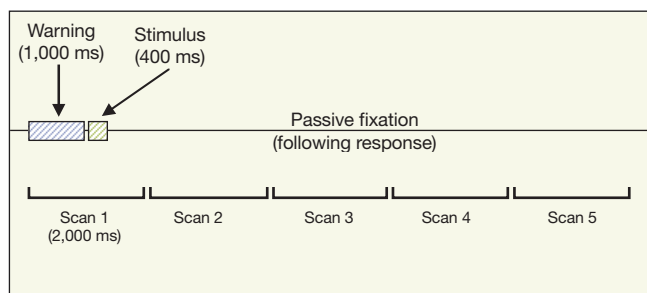


Figure 1 Relative timing of stimulus presentation and scan acquisition.

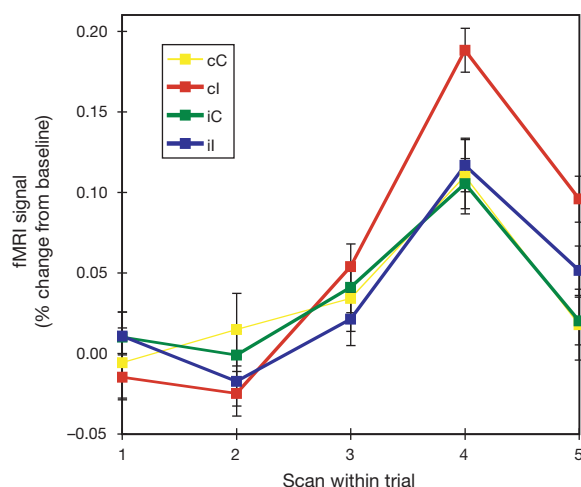


Figure 3 Time-course of activity within the region shown in Fig. 2. Like incompatible trials, compatible trials are displayed according to previous trial-type (cC: previous trial-type compatible; iC: previous trial-type incompatible). fMRI signal is based on the group and area average, expressed as percent change from baseline (average scan-one activation). Error bars are based on the s.e.m. Planned comparisons confirmed that peak activation was higher on incompatible trials than compatible trials ($M = 0.15\%$ and 0.11% , one-tailed $t(7) = 2.32$, $P < 0.03$) and higher for cI trials than iI trials ($M = 0.18\%$ and 0.12% , one-tailed $t(7) = 2.39$, $P < 0.025$).

Functional magnetic resonance imaging (fMRI) data collected as subjects performed the flanker task unambiguously confirmed both predictions of the conflict-monitoring hypothesis. The use of event-related scan acquisition and long interstimulus intervals allowed us to trace the time-course of regional brain activation during the course of individual trials (Fig. 1). As anticipated on the basis of previous findings⁷, activation within the ACC showed a response-related pattern, rising from a baseline at the time of stimulus presentation to a peak about 5 s later (consistent with the 5-s haemodynamic lag typically observed in fMRI experiments)¹⁴. Analyses revealed a region within the dorsal ACC (BA 32) where peak activation was greater during incompatible than compatible trials, and greater during cI than iI trials (Figs 2, 3). No area within the ACC showed the opposite pattern in either comparison.

Several additional findings further support the conflict-monitoring view. First, on compatible trials—unlike incompatible ones—peak ACC activation was not significantly affected by previous trial type (see Fig. 3). This result is consistent with the conflict-monitoring hypothesis, as compatible trials are unlikely to induce conflict, regardless of context. Importantly, when the data for compatible trials is considered alongside that for incompatible ones, it becomes clear that the difference between cI and iI trials is actually part of a larger pattern, involving an interaction between previous and current trial type. A confirmatory analysis (see Methods) showed this interaction to be present in the ACC to a significance level of $P < 0.001$.

Also informative is that, across subjects, the strength of ACC activation was strongly correlated with the severity of conflict. For each subject, we calculated the difference in reaction time between cI and iI trials, using this as a measure of the difference in conflict between these two trial-types. Analysis revealed a robust positive correlation between this index of conflict and the behaviour of the ACC, measured again as the difference between cI and iI trials ($r^2 = 0.66$, $P < 0.01$).

Finally, convergent evidence for the conflict-monitoring theory is provided by the results of a companion study in which the ACC showed comparable behaviour in a different behavioural task (the Stroop task)^{15,16}.

Although our findings challenge the prevailing view of ACC function, note that the conflict-monitoring account does not rule

out an influence of the ACC on selection-for-action. Indeed, behavioural, anatomical and psychophysiological evidence indicates that conflict monitoring may act as a source of feedback to mechanisms involved in recruiting attention, serving to indicate the need for increased top-down control on information processing⁶. An important goal of subsequent studies will be to evaluate this potential relationship between conflict monitoring and cognitive control, as well as to investigate the relation between the ACC and other brain areas implicated in executive function. □

Methods

Subjects and task

Subjects were eleven neurologically normal, right-handed volunteers (seven male, ages 21–32). Stimuli were generated using PsyScope software¹⁷ on a Macintosh computer, and appeared on a back-projection screen mounted inside the scanner bore, which subjects viewed through a mirror. Subjects were instructed to foveate a centrally-located fixation point throughout the task. This point (an asterisk) brightened 1,000 ms before the appearance of the target stimulus. Target stimuli (40% compatible, 60% incompatible in pseudorandom order) appeared just above the fixation point for 400 ms with an interstimulus interval of 10 s. Subjects used their right hand to respond, and were instructed to respond to a left-facing target by pressing a button beneath their index finger, and to a right-facing target by pressing a button beneath their middle finger.

Behavioural data reflected both the basic compatibility effect and the sequential dependency effect reported by Gratton *et al.*¹³. Reaction times were longer for incompatible trials than for compatible ones ($M = 790$ and 523 ms, $t(10) = 4.86$, $P < 0.001$), and the flanker effect (difference between incompatible and compatible conditions) was larger for trials following compatible than incompatible trials ($M = 290$ and 243 ms, $t(10) = 3.97$, $P < 0.005$). cI trials were slower than iI trials for eight of the eleven subjects. Errors made up two percent of trials, and were most common on cI trials, consistent with the Gratton effect and ruling out a speed-accuracy tradeoff.

Image acquisition and analysis

Images were acquired with a 1.5-T GE Signa whole-body scanner with a standard head-coil. Thirteen axial slices (with 3.75^3 cm voxels) were obtained parallel to the AC-PC line. Beginning with brightening of the fixation point, five consecutive 2-s functional scans were acquired on each behavioural trial, using a two-shot spiral T2*-weighted sequence¹⁸, with TR 1,000 ms, TE 35 ms, flip angle 55° and field of view 24 cm. Structural images were obtained with a standard T1-weighted pulse sequence.

Because our second fMRI prediction was intended to apply only in the presence of the Gratton effect, we decided before running the experiment that data would be analysed only for subjects whose behaviour on incompatible trials showed the Gratton effect (cI responses slower than iI responses; eight of our eleven subjects). Only correct trials were included, given the previously observed association of errors with ACC activation⁷. Data from individual participants were subjected to a voxelwise within-block linear detrending and between-block, subtractive mean normalization. Images were co-registered to a common reference structural MRI scan using a 12-parameter automated algorithm¹⁹. Images were then smoothed with an 8-mm full width at half maximum, three-dimensional gaussian filter to accommodate individual differences in anatomy.

To identify areas showing an effect of current trial-type, an analysis of variance (ANOVA) with current trial-type, previous trial-type and scan-within-trial as factors was run on each voxel, and voxels were identified that showed an interaction between scan and current trial-type. To identify areas that were differentially activated during cI and iI trials, we ran a separate ANOVA involving only incompatible trials with scan-within-trial and previous trial-type as factors, and identified voxels that showed an interaction between scan and previous trial-type. In both analyses, we used an eight-voxel cluster-size threshold to correct for multiple comparisons²⁰. Planned comparisons on peak activation were based on average activation across the regions identified by each ANOVA and their intersection. To confirm the presence of an interaction between current and previous trial-types, we performed a two-way ANOVA using the approach above, but focusing on data pooled from scans four and five to maximize sensitivity to task-related effects. This yielded a main effect of current trial-type and an interaction between current and previous trial-types, both appearing in an area of the ACC very similar to that identified in the primary analyses. Both effects survived a significance threshold of $P < 0.005$.

Received 19 April; accepted 23 August 1999.

1. D'Esposito, M. *et al.* The neural basis of the central executive system of working memory. *Nature* **378**, 279–281 (1995).
2. Posner, M. I. & DiGirolamo, G. J. in *The Attentive Brain* (ed. Parasuraman, R.) 401–423 (MIT Press, Cambridge, Massachusetts, 1998).
3. Frith, C. D., Friston, K. J., Liddle, P. F. & Frackowiak, R. S. J. Willed action and the prefrontal cortex in man: A study with PET. *Proc. R. Soc. Lond. B* **244**, 241–246 (1991).
4. Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M. & Raichle, M. E. Positron emission tomography studies of the processing of single words. *J. Cogn. Neurosci.* **1**, 153–170 (1989).
5. Posner, M. I., Petersen, S. E., Fox, P. T. & Raichle, M. E. Localization of operations in the human brain. *Science* **240**, 1627–1631 (1988).
6. Botvinick, M. M., Braver, T. S., Carter, C. S., Barch, D. M. & Cohen, J. D. Evaluating the demand for control: Anterior cingulate cortex and crosstalk monitoring. (Technical Report 98.1, Center for the Neural Basis of Cognition, Pittsburgh, Pennsylvania, 1998).

7. Carter, C. S. *et al.* Anterior cingulate cortex, error detection, and the on-line monitoring of performance. *Science* **280**, 747–749 (1998).
8. Berns, G. S., Cohen, J. D. & Mintun, M. A. Brain regions responsive to novelty in the absence of awareness. *Science* **276**, 1272–1275 (1997).
9. Eriksen, B. A. & Eriksen, C. W. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* **16**, 143–149 (1974).
10. Cohen, J. D., Servan-Schreiber, D. & McClelland, J. L. A parallel distributed processing approach to automaticity. *Am. J. Psychol.* **105**, 239–269 (1992).
11. Coles, M. G. H., Gratton, G., Bashore, T. R., Eriksen, C. W. & Donchin, E. A psychophysiological investigation of the continuous flow model of human information processing. *J. Exp. Psychol. Human Percept. Perform.* **11**, 529–553 (1985).
12. Allport, A. in *Perspectives on Perception and Action* (eds Heuer, H. & Sanders, A. F.) 395–419 (Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1987).
13. Gratton, G., Coles, M. G. H. & Donchin, E. Optimizing the use of information: strategic control of activation and responses. *J. Exp. Psychol. Gen.* **4**, 480–506 (1992).
14. Kwong, K. K. *et al.* Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc. Natl Acad. Sci. USA* **89**, 5675–5679 (1992).
15. Carter, C. S. *et al.* Strategic vs. evaluative contributions of the anterior cingulate cortex to executive processes of cognition. *Proc. Natl Acad. Sci. USA* (submitted).
16. Casey, B. J. *et al.* Ventral and dorsal prefrontal activity as a function of target probability. *J. Cogn. Neurosci.* **10** (Suppl., Annual meeting abstract program), 108 (1998).
17. Cohen, J. D., MacWhinney, B., Flatt, M. & Provost, J. PsyScope: An interactive graphical system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behav. Res. Methods Instr. Comput.* **25**, 257–271 (1993).
18. Noll, D., Cohen, J. D., Meyer, C. H. & Schneider, W. Spiral K-space MR imaging of cortical activation. *J. Magn. Reson. Imag.* **5**, 49–56 (1995).
19. Woods, R. P., Cherry, S. R. & Mazziotta, J. C. Rapid automated algorithm for aligning and reslicing PET images. *J. Comput. Assist. Tomog.* **16**, 620–633 (1992).
20. Forman, S. D. *et al.* Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn. Reson. Med.* **33**, 636–647 (1995).

Acknowledgements

We thank D. Noll for technical assistance. This work was supported by NARSAD, NSF, NIMH and the Fetzer Foundation.

Correspondence and requests for materials should be addressed to M.B. (e-mail: mmb@cncb.cmu.edu).

Two functionally distinct α_2 -adrenergic receptors regulate sympathetic neurotransmission

Lutz Hein*, John D. Altman† & Brian K. Kobilka†

* Department of Pharmacology and Toxicology, University of Würzburg, Würzburg D 97078, Germany

† Howard Hughes Medical Institute and the Departments of Medicine and Molecular and Cellular Physiology, Stanford University, Stanford, California 94305, USA

The sympathetic nervous system regulates cardiovascular function by activating adrenergic receptors in the heart, blood vessels and kidney¹. α_2 -Adrenergic receptors are known to have a critical role in regulating neurotransmitter release from sympathetic nerves and from adrenergic neurons in the central nervous system^{2–5}; however, the individual roles of the three highly homologous α_2 -adrenergic-receptor subtypes (α_{2A} , α_{2B} , α_{2C}) in this process are not known. We have now studied neurotransmitter release in mice in which the genes encoding the three α_2 -adrenergic-receptor subtypes were disrupted. Here we show that both the α_{2A} - and α_{2C} -subtypes are required for normal presynaptic control of transmitter release from sympathetic nerves in the heart and from central noradrenergic neurons. α_{2A} -Adrenergic receptors inhibit transmitter release at high stimulation frequencies, whereas the α_{2C} -subtype modulates neurotransmission at lower levels of nerve activity. Both low- and high-frequency regulation seem to be physiologically important, as mice lacking both α_{2A} - and α_{2C} -receptor subtypes have elevated plasma noradrenaline

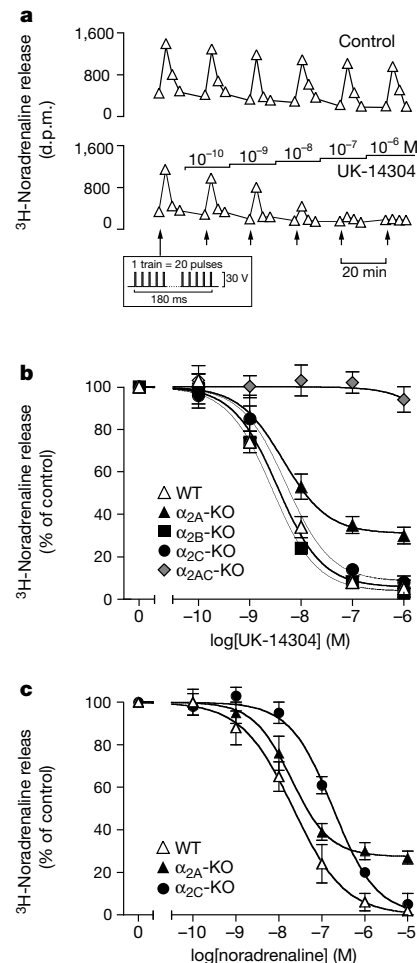


Figure 1 Inhibition of noradrenaline release in atria by α_2 -receptor stimulation. **a**, Atria from wild-type mice (WT) were electrically stimulated (1 train of 20 pulses; 50 Hz) at 20-min intervals. In control conditions, transmitter release was constant over six stimulation periods (upper trace). Increasing concentrations of the α_2 -agonist UK-14304 inhibited the electrically evoked release in WT atria (lower trace). **b**, Concentration–response curves for inhibition by the α_2 -agonist UK-14304 of electrically evoked ^3H -noradrenaline release from isolated atria. **c**, Inhibition of ^3H -noradrenaline release from atria by noradrenaline. Data shown are means $\pm$ s.e.m. from 6–8 atria.

concentrations and develop cardiac hypertrophy with decreased left ventricular contractility by four months of age.

The release of neurotransmitters is modulated by presynaptic autoreceptors that are activated by the released transmitter. Presynaptic inhibitory autoreceptors were first described for noradrenergic neurons and subsequently for other neurotransmitters, including dopamine, acetylcholine, γ -aminobutyric acid, histamine and serotonin^{2–5}. Within the sympathetic nervous system, nine adrenergic receptor subtypes⁶ (three α_1 , α_2 and β -receptors) can be activated by adrenaline and noradrenaline, but only α_2 -adrenergic receptors have been implicated as inhibitory presynaptic autoreceptors. The lack of drugs that are sufficiently selective for the three subtypes has meant that their physiological roles and therapeutic potential have not been fully elucidated. An alternative approach to pharmacological characterization has been the deletion of α_{2B} - or α_{2C} -receptor genes, or *in vivo* site-directed mutagenesis of the α_{2A} -receptor gene by gene targeting in mice^{7–11}. We recently disrupted the gene for the α_{2A} -adrenergic-receptor subtype in mice¹². α_{2A} -receptor-deficient mice (α_{2A} -KO) are produced at the expected mendelian ratios from heterozygote intercrosses. They develop normally and show no gross pathological abnormalities¹².

To investigate the role of the three α_2 -receptor subtypes in presynaptic autoinhibition, we measured ^3H -noradrenaline release

in tissues containing sympathetic (heart atria) or central (brain cortex) noradrenergic axons from normal (WT) and α_2 -receptor-deficient animals. Electrical stimulation with short pulse trains (20 pulses, 50 Hz) elicited the release of ^3H -noradrenaline in the atria of WT and α_{2A} -, α_{2B} - and α_{2C} -KO mice, with no difference in the magnitude of evoked release (Fig. 1). In these conditions, autoinhibition does not affect release because the single pulse trains are too short for the released noradrenaline to activate presynaptic α_2 -receptors^{13,14}. In WT atria, the α_2 -adrenoceptor agonist UK-14304 caused concentration-dependent inhibition of ^3H -noradrenaline release up to $95 \pm 3\%$ (Fig. 1a, b). Similar results were obtained in mice lacking the α_{2B} - or α_{2C} -receptor subtype (Fig. 1b). In mice lacking the α_{2A} -receptor, however, the maximal inhibitory effect of UK-14304 was reduced to $70 \pm 4\%$ (Fig. 1b). We obtained similar results in brain cortex slices, although the effect of the α_2 -agonist in brain cortex from mice lacking the α_{2A} -receptor was smaller than in atria (data not shown). These results indicate that, as has been suggested from pharmacological experiments^{15–18}, the α_{2A} -adrenergic receptor is the principal autoreceptor in the presynaptic feedback loop regulating noradrenaline release. However, the residual effect of UK-14304 in α_{2A} -deficient mice indicates that a second α_2 -autoreceptor operates in both sympathetic neurons and central adrenergic neurons, although the role of this second subtype is much greater in sympathetic neurons.

To determine the identity of the second subtype, we crossed mouse strains lacking α_{2A} -, α_{2B} - or α_{2C} -subtypes to generate mice

lacking two receptor subtypes. The only viable animals lacking two α_2 -receptor subtypes that could be produced from these crosses were mice lacking both α_{2A} - and α_{2C} -receptors (α_{2AC} -KO). In these mice, the inhibitory effect of UK-14304 was completely abolished (Fig. 1b), indicating that the α_{2C} -subtype is also involved in regulating noradrenaline release.

Exogenous noradrenaline inhibited transmitter release dose dependently (Fig. 1c). However, the concentration–response curve for noradrenaline was shifted to the right in atria from α_{2C} -deficient mice compared with WT or α_{2A} -deficient mice (Fig. 1c; EC_{50} (50% excitatory concentration) values: WT, 16 nM; α_{2A} -KO, 20 nM; α_{2C} -KO, 156 nM). This potency difference for noradrenaline may reflect the difference in affinity for noradrenaline between the α_2 -receptor subtypes: noradrenaline has a higher affinity for the α_{2C} -subtype than for the α_{2A} -receptor¹⁹ (K_i inhibitory constant values: α_{2C} , 650 nM; α_{2A} , 5,800 nM). The results of the noradrenaline inhibition studies (Fig. 1c) demonstrate that the α_{2C} -subtype mediates autoinhibition by low concentrations of noradrenaline in wild-type mice. Sensitivity to low catecholamine concentrations may be important for regulating catecholamine release from sympathetic nerves stimulated at low frequencies. To investigate this hypothesis, we stimulated isolated atria at different frequencies to mimic the physiological range of sympathetic nerve activity²⁰. For this experiment, the atria were electrically stimulated with 16 pulse trains and the intervals between the individual trains were varied between 0.5 s (frequency of 2 Hz) and 20 s (frequency of 0.05 Hz). Without presynaptic autoinhibition, one would expect that the amount of transmitter released would be independent of the stimulation frequency. We observed this in α_{2AC} -KO atria (Fig. 2a). However, if presynaptic receptors are inhibiting further transmitter release, the amount of transmitter overflow should decrease at

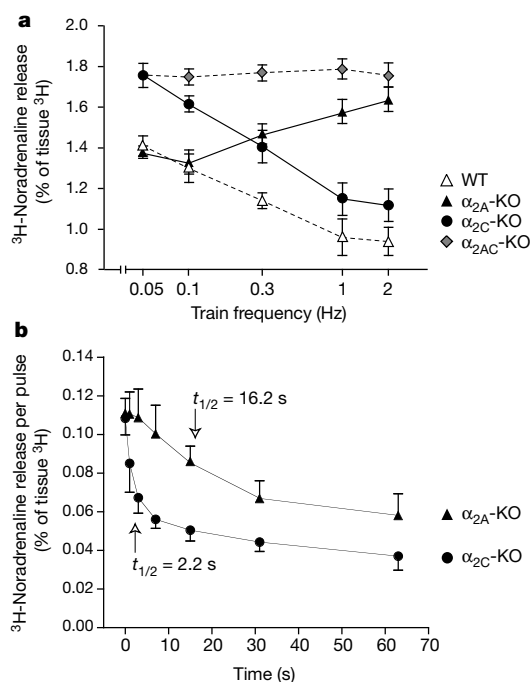


Figure 2 Autoinhibition of noradrenergic neurotransmission in isolated atria is mediated by α_{2A} - and α_{2C} -receptors. **a**, Frequency dependence of noradrenaline release. Isolated atria were stimulated with 16 pulse trains (10 pulses per train at 50 Hz) and the frequency of stimulation was varied between 0.05 and 2 Hz. In wild-type mice, transmitter release decreased with higher stimulation frequencies, whereas release increased in α_{2A} -KO atria. In atria lacking α_{2A} - and α_{2C} -receptors, noradrenaline release was independent of the stimulation frequency. In atria from α_{2C} -deficient mice, the frequency release curve was shifted to the right compared with results for WT mice, indicating that α_{2C} -adrenergic receptors operate to regulate noradrenaline release at low stimulation frequencies. **b**, Time course of inhibition of noradrenaline release from mouse atria by α_{2A} - and α_{2C} -receptors. Isolated atria were stimulated with 1, 2, 3, 4, 8, 16, 32 or 64 pulse trains (10 pulses per train at 50 Hz; train frequency, 1 Hz) and the amount of transmitter released during the first and the following pulse trains was plotted against time. Inhibition mediated by the α_{2A} -subtype was much faster (α_{2C} -KO; $t_{1/2} = 2.2$ s) than inhibition by the α_{2C} -receptor (α_{2A} -KO; $t_{1/2} = 16.2$ s). Data shown are means $\pm$ s.e.m. from 8–10 atria.

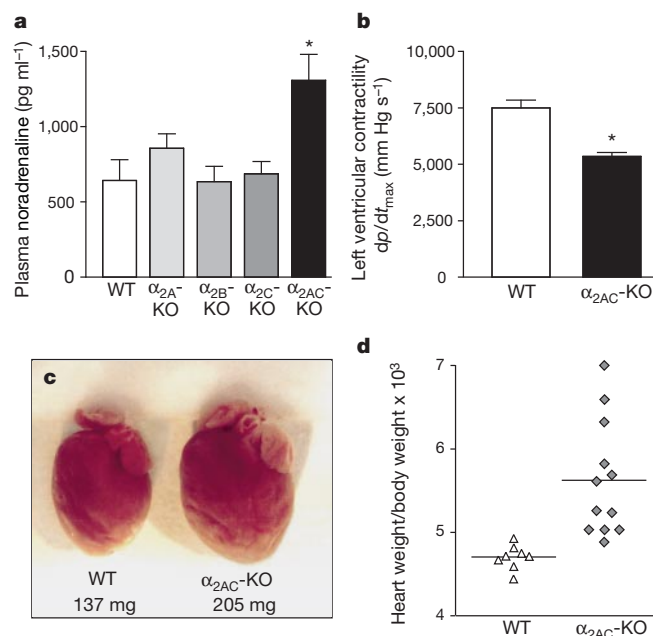


Figure 3 Elevated plasma noradrenaline levels and cardiac phenotype of mice lacking α_2 -adrenergic-receptor subtypes. **a**, Circulating noradrenaline levels were determined in plasma samples obtained from wild-type and α_2 -receptor-deficient mice under anaesthesia. The plasma noradrenaline concentration in α_{2AC} -KO mice was twice as high as in wild-type animals or mice lacking single α_2 -receptor subtypes ($P = 0.0021$, one-way analysis of variance ANOVA followed by Dunnett's tests, $*P < 0.01$). **b**, *In vivo* determination of left ventricular contractility in wild-type and α_{2AC} -KO mice. Left ventricular maximal contractility (dp/dt_{max}) in anaesthetized mice was less in α_{2AC} -KO mice than in wild-type mice. Data are from six animals per group. **c**, Heart of a non-transgenic mouse (WT; left) and an α_{2AC} -KO mouse (right) at 14 weeks of age. Compared with the control heart, the α_{2AC} -KO heart shows significant hypertrophy. **d**, The heart-weight/body-weight ratio was increased in α_{2AC} -KO mice compared with WT mice.

higher synaptic transmitter concentrations, that is, higher stimulation frequencies. Indeed, noradrenaline release from WT atria was significantly less at high stimulation frequencies than at low frequencies (Fig. 2a). In these conditions, this effect can be attributed to presynaptic inhibition of transmitter release by α_{2A} - and α_{2C} -receptors, as noradrenaline overflow from α_{2AC} -KO atria did not change within the frequency range applied. In α_{2A} -deficient atria, noradrenaline release was inhibited at low frequencies, but increased at higher stimulation frequencies; this indicates that the α_{2C} -receptor operates to inhibit noradrenaline release primarily at low frequencies (low noradrenaline concentrations). In contrast, atria from α_{2C} -deficient mice showed no autoinhibition at low stimulation frequencies, confirming that the α_{2C} -receptor inhibits transmitter release at low levels of sympathetic nerve activity and that α_{2A} is required to regulate release at higher frequencies. These results demonstrate that both α_{2A} - and α_{2C} -receptors are required for physiological presynaptic autoinhibition.

The role of the α_{2A} -receptor in mediating autoinhibition at high frequencies (high noradrenaline concentrations) can be attributed to the lower affinity of this subtype for noradrenaline and perhaps a resistance to desensitization. The higher affinity of the α_{2C} -subtype for noradrenaline makes it better suited to respond to the lower concentrations of catecholamines at sympathetic nerve terminals stimulated at low frequencies. To test whether the failure of the α_{2C} -subtype to mediate autoinhibition at high frequencies in this experiment was due to desensitization of this subtype or to slower onset of inhibition, we determined the time course of presynaptic inhibition (Fig. 2b). Transmitter release was elicited with 1 to 64 pulse trains at 1-s intervals and the amount of transmitter released during each pulse was determined. There was no difference in the amount of transmitter released with a single pulse train in atria from α_{2A} -KO or α_{2C} -KO mice. However, the amount of transmitter released during the subsequent pulses declined much faster in atria from α_{2C} -KO mice than in those from α_{2A} -KO mice (Fig. 2b; $t_{1/2}$ (time to half-maximal inhibition) = 2.2 s and 16.2 s, respectively). Thus, coupling of the α_2 -receptor to inhibition of catecholamine release seems to be much faster for the α_{2A} -subtype than for the α_{2C} -subtype. There is no apparent desensitization of either subtype over this time interval.

Several lines of evidence suggest that α_{2A} - and α_{2C} -subtypes are also required for autoinhibition of noradrenaline release *in vivo*. Deletion of the α_{2A} -receptor gene causes tachycardia¹², and increased noradrenaline turnover was observed in mice expressing a mutant α_{2A} -receptor (α_{2A} -D79N)²¹. Thus, we measured plasma noradrenaline levels in mice deficient in α_2 -receptors, and found that, in anaesthetized animals, these levels were dramatically increased in α_{2AC} -KO mice compared with wild-type mice or mice lacking single α_2 -receptor subtypes (Fig. 3a).

To determine the long-term physiological consequences of the altered sympathetic transmitter release in α_2 -receptor-deficient mice, we examined cardiac haemodynamics and the ratios of heart weight to body weight in WT and α_{2AC} -receptor knockout mice. The heart is very sensitive to chronic elevations of catecholamines, and abnormal activity of the sympathetic nervous system has been implicated in the pathogenesis of heart failure^{22–24}. Moreover, cardiac hypertrophy can be induced by chronic infusion of catecholamines in experimental animals. In α_{2AC} -KO mice, left ventricular maximal contractility (dp/dt_{max} ; the maximal change in pressure per unit time) decreased to 70% of the wild-type value at four months of age (Fig. 3b). Mice lacking single α_2 -receptor subtypes showed no impairment of cardiac function at this age (data not shown). In addition, cardiac hypertrophy developed in α_{2AC} -KO mice (Fig. 3c, d). Heart-weight/body-weight ratios were increased in these mice (Fig. 3d), but not in mice lacking only one α_2 -receptor subtype (data not shown). These findings suggest that the heart is exposed to significantly higher levels of catecholamines in α_{2AC} -KO mice than in either α_{2A} -KO or α_{2C} -KO mice. Thus,

inhibition of catecholamine release mediated by both subtypes is physiologically important. □

Methods

Deletion of the α_{2A} -adrenergic-receptor gene

The genomic DNA for the mouse α_{2A} -receptor (*Adra2a*) was cloned from a 129/Sv mouse genomic library (Stratagene). For the gene-targeting vector, a 2.7-kilobase (kb) fragment and an 8-kb fragment were placed around a neomycin-resistance gene, and a herpes simplex virus thymidine kinase gene was added to the 5' end of the plasmid. R1 embryonic stem cells were electroporated and screened for homologous targeting events as described^{12,25}. Chimaeric mice were produced by aggregation of morulae with embryonic stem cells and two chimaeric male mice transmitted the α_{2A} -mutation through the germ line. Homologous recombination at the *Adra2a* locus was identified by Southern blotting using an external genomic probe.

³H-Noradrenaline release from mouse heart atria and brain cortex

This was determined as described^{14,26}, with minor modifications. Briefly, halves of left and right atria or round brain cortex slices (2 mm diameter) were incubated with ³H-noradrenaline (0.1 μ M) in physiological buffer²⁶ or noradrenaline-free medium containing 1 μ M desipramine at 1.8 ml min⁻¹. Transmitter release was elicited by rectangular electrical pulses of 2 ms width and 15 V cm⁻¹ voltage. There were six periods of stimulation in each experiment, applied at 20-min intervals. Drugs were added 15 min before stimulation periods. At the end of the experiments, tissues were solubilized and tritium was determined in superfusate samples and tissues. Electrically evoked overflow of total tritiated compounds was calculated as the difference between total tritium outflow and estimated basal outflow, and was expressed as a percentage of tissue tritium at the time of stimulation. The electrically evoked overflow of total tritium reflects exocytotic release of ³H-noradrenaline²⁷ and is termed thus in this paper. Results are presented as means $\pm$ s.e.m.

Noradrenaline plasma levels

Blood samples were obtained from mice anaesthetized with tribromoethanol (13 μ l of 2.5% solution per g body weight) and placed on a 37 °C table. After 20-min anaesthesia, 400 μ l blood was drawn from the left carotid artery and centrifuged. Plasma noradrenaline was quantified by HPLC combined with electrochemical detection²⁸. Results are displayed as means $\pm$ s.e.m. from 8–12 mice.

Cardiac haemodynamics

Mice were anaesthetized with tribromoethanol and catheterized with a 1.8 French high-fidelity catheter-tip micromanometer (Millar Instruments) as described²⁹.

Cardiac pathology

Hearts from WT and KO animals were fixed in 4% paraformaldehyde in PBS as described³⁰. Mid-ventricular cross-sections were used for computerized morphometric analysis of myocyte cross-sectional areas³⁰. For determination of heart weights, the ventricles of 14-week-old animals were separated from the atria and surrounding connective tissue, rinsed in physiological buffer and briefly blotted dry. Heart weight was normalized to body weight and tibia length.

Received 23 June; accepted 2 September 1999.

1. Taylor, A. A. Autonomic control of cardiovascular function: clinical evaluation in health and disease. *J. Clin. Pharmacol.* **34**, 363–374 (1994).
2. Miller, R. J. Presynaptic receptors. *Annu. Rev. Pharmacol. Toxicol.* **38**, 201–227 (1998).
3. Langer, S. Z. 25 years since the discovery of presynaptic receptors: present knowledge and future perspectives. *Trends Pharmacol. Sci.* **18**, 95–99 (1997).
4. Starke, K., Göthert, M. & Kilbinger, H. Modulation of neurotransmitter release by presynaptic autoreceptors. *Physiol. Rev.* **69**, 864–989 (1989).
5. Stjärne, L. Basic mechanisms and local modulation of nerve impulse-induced secretion of neurotransmitters from individual sympathetic nerve varicosities. *Rev. Physiol. Biochem. Pharmacol.* **112**, 4–137 (1989).
6. Bylund, D. B. *et al.* International Union of Pharmacology nomenclature of adrenoceptors. *Pharmacol. Rev.* **46**, 121–136 (1994).
7. Link, R. E. *et al.* Targeted inactivation of the gene encoding the mouse α_{2C} -adrenoceptor homolog. *Molec. Pharmacol.* **48**, 48–55 (1995).
8. Link, R. E. *et al.* Cardiovascular regulation in mice lacking α_2 -adrenergic receptor subtypes b and c. *Science* **273**, 803–805 (1996).
9. MacMillan, L. B., Hein, L., Smith, M. S., Piascik, M. T. & Limbird, L. E. Central hypotensive effects of the α_{2A} -adrenergic receptor subtype. *Science* **273**, 801–803 (1996).
10. Sallinen, J. *et al.* Genetic alteration of α_{2C} -adrenoceptor expression in mice: influence on locomotor, hypothermic, and neurochemical effects of dexmedetomidine, a subtype-nonspecific α_2 -adrenoceptor agonist. *Molec. Pharmacol.* **51**, 36–46 (1997).
11. Sallinen, J., Haapalinna, A., Vitamaa, T., Kobilka, B. K. & Scheinin, M. Adrenergic α_{2C} -receptors modulate the acoustic startle reflex, prepulse inhibition, and aggression in mice. *J. Neurosci.* **18**, 3035–3042 (1998).
12. Altman, J. D. *et al.* Abnormal regulation of the sympathetic nervous system in α_{2A} -adrenergic receptor knockout mice. *Molec. Pharmacol.* **56**, 154–161 (1999).
13. Singer, E. A. Transmitter release from brain slices elicited by single pulses: a powerful method to study presynaptic mechanisms. *Trends Pharmacol. Sci.* **9**, 274–276 (1988).

14. Wahl, C. A., Trendelenburg, A. U. & Starke, K. Presynaptic α_2 -autoreceptors in mouse heart atria: evidence for the α_{2D} subtype. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **354**, 253–261 (1996).
15. Trendelenburg, A. U., Limberger, N. & Starke, K. Presynaptic α_2 -autoreceptors in brain cortex: α_{2D} in the rat and α_{2A} in the rabbit. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **348**, 35–45 (1993).
16. Trendelenburg, A. U. *et al.* A re-investigation of questionable subclassifications of presynaptic α_2 -autoreceptors: rat vena cava, rat atria, human kidney and guinea-pig urethra. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **356**, 721–737 (1997).
17. Molderings, G. J. & Göthert, M. Subtype determination of presynaptic α_2 -autoreceptors in the rabbit pulmonary artery and human saphenous vein. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **352**, 483–490 (1995).
18. Smith, K., Gavin, K. & Docherty, J. R. Investigation of the subtype of α_2 -adrenoceptor mediating prejunctional inhibition of cardioacceleration in the pithed rat heart. *Br. J. Pharmacol.* **115**, 316–320 (1995).
19. Link, R. E., Daunt, D., Barsh, G., Chruscinski, A. & Kobilka, B. Cloning of two mouse genes encoding α_2 -adrenergic receptor subtypes and identification of a single amino acid in the mouse α_2 -C10 homolog responsible for an interspecies variation in antagonist binding. *Molec. Pharmacol.* **42**, 16–27 (1992).
20. McLachlan, E. M., Davies, P. J., Habler, H. J. & Jamieson, J. On-going and reflex synaptic events in rat superior cervical ganglion cells. *J. Physiol. (Lond.)* **501**, 165–181 (1997).
21. Lakhani, P. P. *et al.* Substitution of a mutant α_{2A} -adrenergic receptor via "hit and run" gene targeting reveals the role of this subtype in sedative, analgesic, and anesthetic-sparing responses *in vivo*. *Proc. Natl Acad. Sci. USA* **94**, 9950–9955 (1997).
22. Packer, M. The neurohormonal hypothesis: a theory to explain the mechanism of disease progression in heart failure. *J. Am. Coll. Cardiol.* **20**, 248–254 (1992).
23. Cohn, J. N. *et al.* Plasma norepinephrine as a guide to prognosis in patients with chronic congestive heart failure. *N. Engl. J. Med.* **311**, 819–823 (1984).
24. Lechat, P. *et al.* Clinical effects of β -adrenergic blockade in chronic heart failure: a meta-analysis of double-blind, placebo-controlled, randomized trials. *Circulation* **98**, 1184–1191 (1998).
25. Hein, L., Barsh, G. S., Pratt, R. E., Dzau, V. J. & Kobilka, B. K. Behavioural and cardiovascular effects of disrupting the angiotensin II type-2 receptor in mice. *Nature* **377**, 744–747 (1995).
26. Limberger, N., Trendelenburg, A. A. & Starke, K. Subclassification of presynaptic α_2 -adrenoceptors: α_{2D} -autoreceptors in mouse brain. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **352**, 43–48 (1995).
27. Taube, H. D., Starke, K. & Borowski, E. Presynaptic receptor systems on the noradrenergic neurones of rat brain. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **299**, 123–141 (1977).
28. Graefe, K. H. *et al.* 1,1'-Diisopropyl-2,4'-cyanine (disprocynium24), a potent uptake2 blocker, inhibits the renal excretion of catecholamines. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **356**, 115–125 (1997).
29. Engelhardt, S., Hein, L., Wiesmann, F. & Lohse, M. J. progressive hypertrophy and heart failure in β_1 -adrenergic receptor transgenic mice. *Proc. Natl Acad. Sci. USA* **96**, 7059–7064 (1999).
30. Hein, L. *et al.* Overexpression of angiotensin AT1 receptor transgene in the mouse myocardium produces a lethal phenotype associated with myocyte hyperplasia and heart block. *Proc. Natl Acad. Sci. USA* **94**, 6391–6396 (1997).

Acknowledgements

This work was supported by grants from the Deutsche Forschungsgemeinschaft SFB355 (to L.H.) and the Howard Hughes Medical Institute (to B.K.). We thank K. Starke, A. U. Trendelenburg, M. J. Lohse and M. M. Bücheler for consultation and assistance with transmitter release experiments; S. Engelhardt for assistance with the pathological analysis; R. Link for help with construction of the α_{2A} -targeting vector; A. Chruscinski for help in characterizing the α_{2AC} -KO mice; and K. H. Graefe and R. Wölfl for HPLC analysis of plasma noradrenaline levels.

Correspondence and requests for materials should be addressed to B.K.K., Howard Hughes Medical Institute, B157 Beckman Center, Stanford University, Stanford, CA 94305 (e-mail: kobilka@cmgm.stanford.edu).

Rapid gating and anion permeability of an intracellular aquaporin

Masato Yasui[†], Akihiro Hazama^{†,‡}, Tae-Hwan Kwon[§], Søren Nielsen[§], Wm. B. Guggino[‡] & Peter Agre^{*}

Departments of ^{*}Biological Chemistry, Medicine, and [‡]Physiology, Johns Hopkins University School of Medicine, Baltimore MD 21205, USA

[§]Department of Cell Biology, Institute of Anatomy, University of Aarhus, DK-8000, Denmark

[†] These authors contributed equally to this work

Aquaporin (AQP) water-channel proteins are freely permeated by water but not by ions or charged solutes¹. Although mammalian aquaporins were believed to be located in plasma membranes, rat AQP6 is restricted to intracellular vesicles in renal epithelia². Here we show that AQP6 is functionally distinct from other known

aquaporins. When expressed in *Xenopus laevis* oocytes, AQP6 exhibits low basal water permeability; however, when treated with the known water channel inhibitor, Hg²⁺, the water permeability of AQP6 oocytes rapidly rises up to tenfold and is accompanied by ion conductance. AQP6 colocalizes with H⁺-ATPase in intracellular vesicles of acid-secreting α -intercalated cells in renal collecting duct. At pH less than 5.5, anion conductance is rapidly and reversibly activated in AQP6 oocytes. Site-directed mutation of lysine to glutamate at position 72 in the cytoplasmic mouth of the pore changes the cation/anion selectivity, but leaves low pH activation intact. Our results demonstrate unusual biophysical properties of an aquaporin, and indicate that anion-channel function may now be explored in a protein with known structure.

The function of AQP6 was previously unknown. The amino-acid sequence of AQP6 (ref. 3) is most closely related to those of AQP2 from principal cells of renal collecting duct⁴ and AQP0 (major intrinsic protein) from lens⁵. Unlike AQP2 (ref. 6), AQP6 does not travel to the plasma membrane². Like AQP0 (ref. 7), AQP6 has been reported to have low water permeability⁸. When compared with water-injected *X. laevis* oocytes (water permeability (P_f) = $5.3 \pm 0.3 \text{ cm s}^{-1} \times 10^{-4}$), osmotic water permeability of oocytes expressing rat AQP6 is only slightly increased (P_f = $12.0 \pm 0.9 \text{ cm s}^{-1} \times 10^{-4}$) (Fig. 1a).

Mercury ions inhibit the water permeability of most aquaporins¹; HgCl₂ was also reported to inhibit AQP6 (ref. 8), but, unexpectedly, we found that oocytes expressing rat AQP6 are activated by Hg²⁺ (Fig. 1a). Within seconds, exposure to 300 μM HgCl₂ markedly increased the osmotic water permeability of AQP6 oocytes (P_f = $93.0 \pm 4.3 \text{ cm s}^{-1} \times 10^{-4}$), which was reversed with 5 mM β -mercaptoethanol (P_f = $20.0 \pm 3.9 \text{ cm s}^{-1} \times 10^{-4}$). Other aquaporins have low activation energies, $E_A < 4 \text{ kcal mol}^{-1}$, which rise after treatment with Hg²⁺ (ref. 9). For AQP6, E_A shifted from 12.5 to 3.6 kcal mol⁻¹ when the oocytes were preincubated with HgCl₂ (Fig. 1b). Hg²⁺ could not have activated water permeability by inducing vesicle trafficking to the cell surface, because

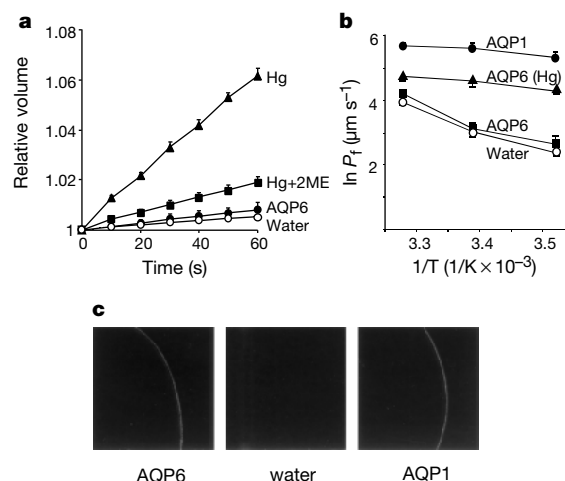


Figure 1 Water permeability of AQP6. **a**, Time-dependent osmotic swelling of water-injected oocytes (open circles), AQP6 oocytes without treatment (filled circles), AQP6 oocytes incubated with 300 μM HgCl₂ (triangles), and oocytes incubated with HgCl₂ followed by 5 mM β -mercaptoethanol (squares). **b**, Arrhenius activation energy (E_A) of osmotic water permeabilities of water-injected oocytes (open circles), untreated AQP1 oocytes (filled circles), untreated AQP6 oocytes (squares) and AQP6 oocytes incubated with 300 μM HgCl₂ (triangles). Osmotic-swelling assays were then performed at 11, 22 or 32 °C (means $\pm$ s.e., $n = 5$). **c**, Confocal microscopy of oocytes injected with AQP6 cRNA, water or AQP1 cRNA after fixation, and incubation with anti-AQP6 (left and middle panels) or anti-AQP1 (right panel).

AQP6 is located in the oocyte plasma membrane (Fig. 1c), where it appears to be directly gated.

We evaluated the possible ionic permeation of AQP6 oocytes using a two-electrode voltage clamp⁹. Application of 100 μ M HgCl₂ rapidly induced a large increase in membrane current that was reversed with 5 mM β -mercaptoethanol (Fig. 2a, b). The current was not time or voltage-dependent under our experimental conditions (Fig. 2a, left). Water-injected oocytes (Fig. 2a, right), and AQP0, AQP1 and AQP2 oocytes failed to exhibit inducible currents (data not shown). The slope of the Hg²⁺-stimulated current voltage relation for AQP6 oocytes was almost linear with a reversal potential of -22.1 ± 3.8 mV in 100 mM NaCl solution (Fig. 2b), and the HgCl₂ effect was dose dependent between 1 μ M and 100 μ M (Fig. 2c). Incubation of AQP6 oocytes in cAMP, phorbol dibutyrate or ATP did not affect the water permeability or ion conductance (data not shown).

Site-directed mutagenesis was undertaken to identify the residues responsible for Hg²⁺ activation of AQP6. There are three cysteines in the AQP6 molecule, but wild-type AQP6 and the C6S mutant exhibited similar Hg²⁺-inducible water permeability. Although located in the plasma membrane, C155A and C190A each exhibited 50% less Hg²⁺-inducible water permeability and ion conductance, and the double mutant C155A–C190A exhibited negligible Hg²⁺ activation (data not shown). Thus, Hg²⁺ appears to trap AQP6 in a conformational state that is permeable to water and ions.

AQP6 is abundant in intracellular vesicles of acid-secreting α -intercalated cells of renal collecting duct², where immunogold electron microscopy showed the colocalization of AQP6 and H⁺-ATPase (Fig. 3). This indicated that low pH may be the natural activator of AQP6. We found that the water permeability of AQP6

oocytes rises slightly: $P_f = 12.2 \pm 1.5$ cm s $\times 10^{-4}$ (at pH 7.5); $P_f = 32.2 \pm 1.1$ cm s $\times 10^{-4}$ (at pH 4.0). In contrast, the osmotic water permeability of water-injected oocytes was not appreciably altered by low pH: $P_f = 7.4 \pm 0.7$ cm s $\times 10^{-4}$ (at pH 7.5); $P_f = 9.0 \pm 0.6$ cm s $\times 10^{-4}$ (at pH 4.0) (Fig. 4a). A membrane current rapidly appeared in AQP6 oocytes at pH 4.0, but disappeared after return to pH 7.5 (Fig. 4b, c). These membrane currents were time independent and slightly outward rectifying (Fig. 4b, c). In contrast, water-injected oocytes (Fig. 4b) or oocytes expressing AQP0, AQP1, AQP2, AQP4 or AQP5 (data not shown) failed to exhibit increased ion conductance in response to low pH. The ion conductance of AQP6 oocytes was observed at pH 5.5, and increased markedly at pH 4.0 (Fig. 4d), but could not be reliably measured at lower pHs in which oocytes become unstable. Attempts to study membrane currents in transiently transfected mammalian cell lines MDCK, HEK 293 and CHO were prevented by restriction of AQP6 to intracellular sites (data not shown).

Multiple studies indicated that the membrane currents in AQP6 oocytes are inherent channel permeation properties of AQP6 rather than endogenous channels. When the extracellular NaCl concentration was reduced to from 100 to 10 mM, the reversal potential of AQP6 oocytes at pH 4.0 shifted (from -27.5 ± 1.6 mV to $+3.5 \pm 6.3$ mV; Fig. 5a, left), indicating that the low pH induced current is much more selective to Cl[−] ions than to Na⁺ ions ($P_{Na}/P_{Cl} = 0.28 \pm 0.09$). Thus, AQP6 is less selective than the chloride ClC-1 channel (ratio 0.10; ref. 10) but more selective than voltage-dependent anion channels (VDACs) (ratio 0.59; ref. 11). Selectivity of ion channels is often determined by specific charged residues^{10,12,13}, and in the 'hourglass' structural model for aquaporins¹⁴, the unique AQP6 residue K72 is positioned at the cytoplasmic mouth of the aqueous pore (Fig. 5d). This charge was reversed by site-directed mutagenesis to K72E. In contrast to wild-type AQP6 oocytes, K72E oocytes exhibited similar reversal potential at pH 4.0 when measured in 100 mM NaCl (-16.6 ± 1.1 mV, $n = 4$) or 10 mM NaCl (-17.4 ± 12.1 mV, $n = 3$) (Fig. 5a, right), indicating that cation permeability had risen to an equivalent level in the mutant ($P_{Na}/P_{Cl} = 1.13 \pm 0.52$). The mutant K72H did not exhibit altered ion selectivity (data not shown).

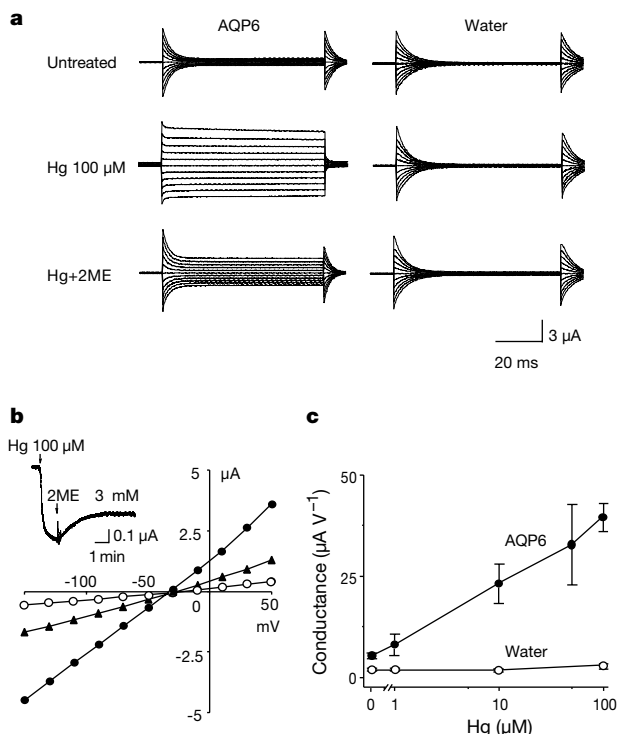


Figure 2 Electrophysiological analysis of AQP6. **a**, Representative currents of AQP6 oocytes (left) and water-injected oocytes (right) without treatment (upper panels), 1 min after the application of 100 μ M HgCl₂ (middle panels), or 15 min after the application of 5 mM β -mercaptoethanol to HgCl₂-treated oocytes (lower panels). **b**, Representative I - V curves of AQP6 oocyte in NaCl solution without treatment (open circles), after 100 μ M HgCl₂ (filled circles) or after HgCl₂ followed by 5 mM β -mercaptoethanol (triangles). AQP6 oocyte membrane currents were continuously measured in at -50 mV (inset). **c**, Dose-dependent conductance of AQP6 oocytes (filled circles) and water-injected oocytes (open circles) treated with HgCl₂ (means $\pm$ s.e., $n = 4$).

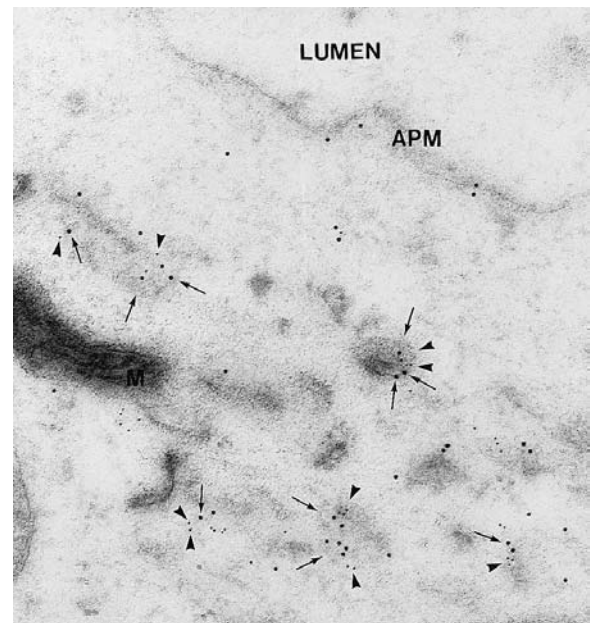


Figure 3 Immunogold electron microscopic colocalization of AQP6 (5-nm particles, arrowheads) and H⁺-ATPase (10-nm particles, arrows) in intracellular vesicles of α -intercalated cells of rat renal collecting duct. Note that some H⁺-ATPase is also observed in the apical plasma membrane (APM) facing the tubular lumen. Magnification 57,000 $\times$.

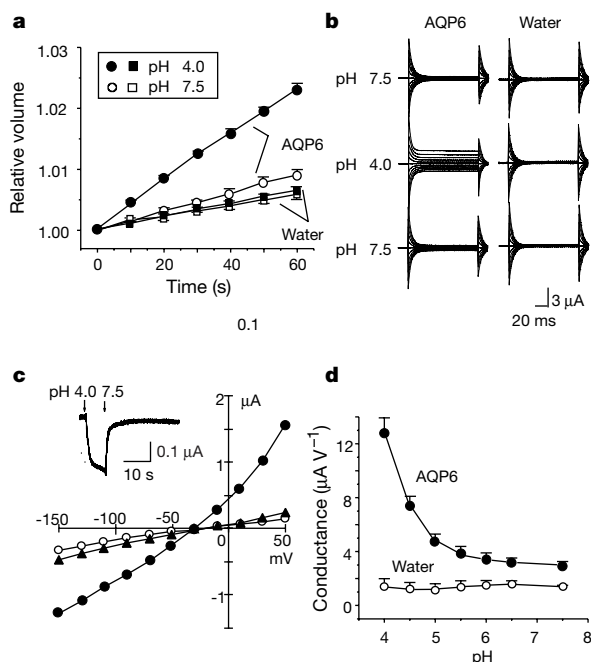


Figure 4 Effect of pH on water permeability and ion conductance of AQP6. **a**, Time-dependent osmotic swelling of AQP6 oocytes (circles) and water-injected oocytes (squares) at pH 7.5 (open symbols) and pH 4.0 (filled symbols). **b**, Representative currents of AQP6 oocytes (left) and water-injected oocytes (right) at pH 7.5 (upper panels), after 1 min at pH 4.0 (middle panels), or 1 min after return to pH 7.5 (lower panels). **c**, Representative I - V curves of AQP6 oocyte at pH 7.5 (open circles), after 1 min at pH 4.0 (filled circles), or 1 min after return to pH 7.5 (triangles). AQP6 oocyte membrane currents were continuously measured at -50 mV (inset). **d**, Relationship between ion conductance and extracellular pH obtained from AQP6 oocytes (filled circles) or water-injected oocytes (open circles) (means $\pm$ s.e., $n = 4$).

Ion selectivity was further evaluated by replacement experiments in which equimolar replacement of 100 mM NaCl with sodium gluconate yielded a more positive reversal potential in AQP6 oocytes at pH 4.0 ($+13.7 \pm 5.8$ mV, $n = 4$; Fig. 5b, left). In contrast, no significant shift was observed when K72E mutant oocytes were studied in 100 mM NaCl (see above) or 100 mM sodium gluconate at pH 4.0 (-6.9 ± 5.6 mV, $n = 4$; Fig. 5b, right). Attempts to modify AQP6 ion current with known chloride channel blockers (DIDS, NPPB, niflumic acid and DPC) or metals (cadmium and zinc) did not result in qualitative alterations (data not shown). Equimolar replacement of 100 mM NaCl with *N*-methyl-D-glucamine (NMDG) hydrochloride did not alter reversal potential in AQP6 oocytes or K72E mutant oocytes at pH 4.0 (Fig. 5c), indicating that NMDG⁺ is as permeable as Na⁺ in wild-type AQP6 and K72E, although both cations are much less permeable than Cl⁻ in wild type.

Our studies establish several unusual biophysical properties for an aquaporin. Water permeability of AQP6 is pharmacologically activated by Hg²⁺ and mimics the proposed physiological gating by acid pH. The anion permeation of AQP6 is particularly provocative, as the 3D structure of aquaporins is known at 4.5–6 Å resolution^{15,16}. The possibility that this conductance is due to an associated chloride channel must be small, as similar currents were not observed in any of our control studies. Although these studies were performed with rat AQP6, we have obtained qualitatively similar results with human AQP6 (data not shown). Moreover, using a classic approach^{10,12,13}, the positively charged pore-lining residue K72 was shown to be critical to selectivity. Although the overall anion selectivity exhibited by AQP6 is less stringent than other chloride channels, the restriction of AQP6 to specific intracellular sites implicates the protein in specific biological processes. Intracellular organelles are acidified by the V-type H⁺-ATPase, and

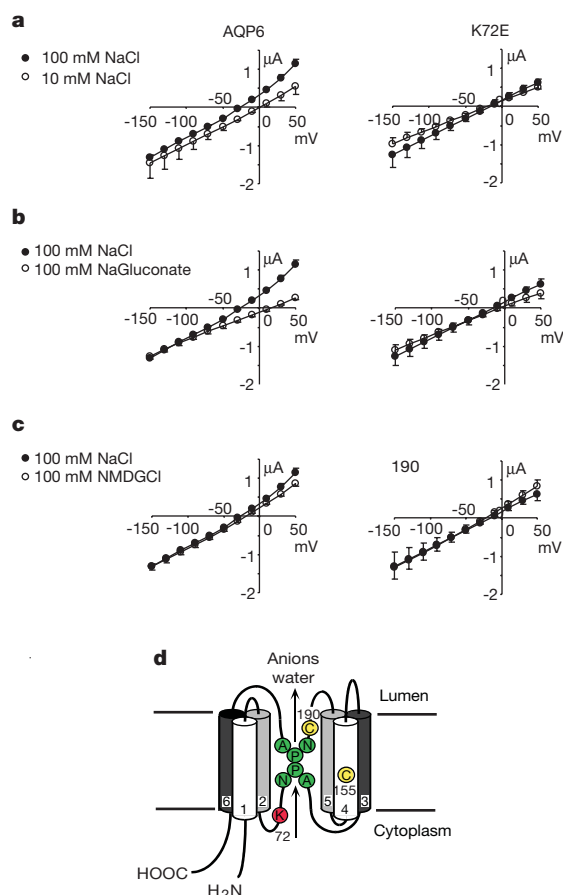


Figure 5 Electrophysiology of wild-type AQP6 and the K72E mutant. **a–c**, Representative I - V curves obtained at pH 4.0 from oocytes expressing wild-type AQP6 (left) or K72E mutant (right) in 100 mM NaCl (filled circles) and 10 mM NaCl solution (open circles) with osmolarity preserved by addition of sucrose (**a**), 100 mM sodium gluconate (**b**) and 100 mM NMDGCl (**c**) ($n = 3–7$). **d**, Transmembrane topology of the AQP6 channel pore. Highly conserved N, P and A residues in internal and external loops that form the aqueous pore are shown in green; pore-lining charged residue K72 is shown in red; and Hg²⁺ activation sites C155 and C190 are shown in yellow. Vesicle lumen and cytoplasmic space are indicated.

vesicles within acid-secreting α -intercalated cells of renal collecting duct are pH 5.0 or lower¹⁷. Anion conductance must occur in acidic vesicles to maintain electroneutrality¹⁸, and the intracellular chloride channel ClC-5 colocalizes with V-type H⁺-ATPase in kidney¹⁹; however, ClC-5 is inhibited at pH less than 6.5 (ref. 12). Thus, the AQP6 conductance is apparently turned on at low pH when ClC-5 is turned off, indicating that ClC-5 may function at an early stage of acidification and AQP6 at a later stage. The water permeability of activated AQP6 may also be important, because in intracellular vesicles AQP6 may contribute to vesicle swelling and membrane fusion during exocytosis or other cellular processes²⁰. Thus, the biological repertoire of aquaporins may be considerably larger than previously believed. □

Methods

Expression in oocytes and measurement of P_f

cRNA was prepared from full-length rat AQP6 cDNA (Gene Bank accession code AF083879) in *Xenopus* oocyte expression construct pXBG (ref. 9). Mutants were introduced with the Chameleon double-stranded, site-directed mutagenesis kit (Stratagene). Defolliculated stage V–VI *Xenopus laevis* oocytes were injected with 50 nl of water or 5–10 ng of cRNA, and swelling was measured directly after transfer from 200 mosM to 70 mosM modified Barth's buffer at 22 °C (ref. 9). Seasonal variations in oocyte quality affected AQP6 expression and water permeability measurements. For Hg²⁺ studies, oocytes were incubated for 5 min in MBS containing 300 μ M HgCl₂ before osmotic swelling, or were incubated with HgCl₂ and then in 5 mM DME for 15 min before

osmotic swelling. For pH studies, oocytes were incubated for 1 min in 200 mosM modified Barth's buffer containing 5 mM HEPES pH 7.5 or 10 mM MES pH 4.0 before osmotic swelling. Immunolocalization of AQP6, AQP6 mutants and AQP1 in oocytes were done as described¹³ using fluorescein-isothiocyanate-conjugated second antibody.

Electrophysiology

Recordings were performed at 22 °C with iso-osmotic NaCl solution (100 mM NaCl, 2 mM KCl, 1 mM MgCl₂, and 5 mM Hepes pH 7.5 or 10 mM MES pH 4–6.5). Oocyte membrane potentials were controlled by two-microelectrode voltage clamp (Axoclamp-2A and pCLAMP software, Axon instruments). These recording conditions permitted analysis of oocyte steady-state currents rather than initial capacitive deflections. Current signals were sampled at 100 µsec, and in most experiments the membrane potential was held at $V_{\text{hold}} = -50$ mV and rapidly stepped to values between +50 and -150 mV. Pulse durations of 100 ms and currents from 10 runs were averaged to reduce noise. Conductances were calculated from slopes of $I-V$ curves between +10 mV and -90 mV. Only healthy oocytes with resting potentials of -25 to -40 mV were used. High concentrations of HgCl₂ (>1 mM) induced background conductance in all oocytes including water-injected oocytes and AQP1 oocytes; this endogenous conductance was partially blocked by 1 mM cadmium. $P_{\text{Na}}/P_{\text{Cl}}$ was calculated using the Goldman-Hodgkin-Katz equation assuming intracellular [Cl⁻] and -[Na⁺] are 30 mM. Although forskolin-induced cation conductance in AQP1 oocytes has been reported²², it was not confirmed by multiple laboratories²³ or during measurements undertaken during these studies.

Immunoelectron microscopy

Sections were prepared from the inner stripe of the outer medulla of Wistar rat kidney embedded by cryosubstitution in Lowicryl HM20 (ref. 2). Double immunogold labelling was performed with affinity-purified polyclonal rabbit anti-AQP6 (ref. 2) and an antibody raised to the carboxy terminus of B1 subunit of H⁺-ATPase²¹. Sequential labellings were performed with blocking steps using rabbit serum and goat-anti-rabbit FAB. Visualization was with goat-anti-rabbit IgG conjugated to 5- or 10-nm colloidal gold particles. Control labellings used affinity-purified rabbit anti-AQP2; absence of background staining was shown by the double labelling procedure (data not shown).

Received 28 July; accepted 21 September 1999.

- Agre, P., Bonhivers, M. & Borgnia, M. J. The aquaporins, blueprints for cellular plumbing systems. *J. Biol. Chem.* **273**, 14659–14662 (1998).
- Yasui, M., Kwon, T.-H., Knepper, M. A., Nielsen, S. & Agre, P. Aquaporin-6: an intracellular vesicle water channel protein in renal epithelia. *Proc. Natl Acad. Sci. USA* **96**, 5808–5813 (1999).
- Ma, T., Frigeri, A., Skach, W. & Verkman, A. S. Cloning of a novel rat kidney cDNA homologous to CHIP28 and WCH-CD water channels. *Biochem. Biophys. Res. Commun.* **197**, 654–659 (1993).
- Fushimi, K. *et al.* Cloning and expression of apical membrane water channel of rat kidney collecting tubule. *Nature* **361**, 549–552 (1993).
- Gorin, M. B., Yancey, S. B., Cline, J., Revel, J. P. & Horwitz, J. The major intrinsic protein (MIP) on the bovine lens fiber membrane. *Cell* **39**, 49–59 (1984).
- Nielsen, S. *et al.* Vasopressin increases water permeability of kidney collecting duct by inducing translocation of aquaporin-CD water channels to plasma membrane. *Proc. Natl Acad. Sci. USA* **92**, 1013–1017 (1995).
- Mulders, S. M. *et al.* Water channel properties of major intrinsic protein of lens. *J. Biol. Chem.* **270**, 9010–9016 (1995).
- Ma, T., Yang, B., Kuo, W. L. & Verkman, A. S. cDNA cloning and gene structure of a novel water channel expressed exclusively in human kidney: evidence for a gene cluster of aquaporins at chromosome locus 12q13. *Genomics* **35**, 543–550 (1996).
- Preston, G. M., Carroll, T. P., Guggino, W. B. & Agre, P. Appearance of water channels in *Xenopus* oocytes expressing red cell CHIP28 protein. *Science* **256**, 385–387 (1992).
- Fahlke, C., Yu, H. T., Beck, C. L., Rhodes, T. H. & George, A. L. Jr Pore-forming segments in voltage-gated chloride channels. *Nature* **390**, 529–532 (1997).
- Roos, N., Benz, R. & Brdiczka, D. Identification and characterization of the pore-forming protein in the outer membrane of rat liver mitochondria. *Biochem. Biophys. Acta* **686**, 204–214 (1982).
- Friedrich, T., Breiderhoff, T. & Jentsch, T. J. Mutational analysis demonstrates that CIC-4 and CIC-5 directly mediate plasma membrane currents. *J. Biol. Chem.* **274**, 896–902 (1999).
- Blachly-Dyson, E. *et al.* Selectivity changes in site-directed mutants of the VDAC ion channel: structural implications. *Science* **247**, 1233–1236 (1990).
- Jung, J. S., Preston, G. M., Smith, B. L., Guggino, W. B. & Agre, P. Molecular structure of the water channel through aquaporin CHIP. The hourglass model. *J. Biol. Chem.* **269**, 14648–14654 (1994).
- Walz, T. *et al.* The three-dimensional structure of aquaporin-1. *Nature* **387**, 624–627 (1997).
- Mitsuoka, K. *et al.* Short α -helices in hourglass pore-forming domains of aquaporin-1, a water channel protein, visualized at 4.5 Å resolution. *J. struct. Biol.* (in the press).
- Schwartz, G. J. & Al-Awqati, Q. Regulation of transepithelial H⁺ transport by exocytosis and endocytosis. *Annu. Rev. Physiol.* **48**, 153–161 (1986).
- Glickman, J., Croen, K., Kelly, S. & Al-Awqati, Q. Golgi membranes contain an electrogenic H⁺ pump in parallel to a chloride conductance. *J. Cell Biol.* **97**, 1303–1308 (1983).
- Günther, W., Luchow, A., Cluzeaud, F., Vandewalle, A. & Jentsch, T. J. CIC-5, the chloride channel mutated in Dent's disease, colocalizes with the proton pump in endocytotically active kidney cells. *Proc. Natl Acad. Sci. USA* **95**, 8075–8080 (1998).
- Jena, B. P. *et al.* Gi regulation of secretory vesicle swelling examined by atomic force microscopy. *Proc. Natl Acad. Sci. USA* **94**, 13317–13322 (1997).
- Hirsch, S. *et al.* Isolation and sequence of a cDNA clone encoding the 31-kDa subunit of bovine kidney vacuolar H⁺-ATPase. *Proc. Natl Acad. Sci. USA* **85**, 3004–3008 (1988).
- Yool, A. J., Stamer, W. D. & Regan, J. W. Forskolin stimulation of water and cation permeability in aquaporin-1 water channels. *Science* **273**, 1216–1218 (1996).
- Agre, P. *et al.* Aquaporins and ion conductance. *Science* **275**, 1490–1492 (1997).

Acknowledgements

We thank M. Lu for the AQP6 plasmid, M. A. Knepper for the anti-H⁺-ATPase, and M. R. T. Hall for confocal microscopy. Support was provided by grants from the NIH and the Cystic Fibrosis Foundation (P.A. and W.B.G.), the Human Frontier Science Program (M.Y.), and the Novo Nordic Foundation, the Karen Elise Jensen Foundation, the Danish Medical Research Council, the Biomembrane Research Center at University of Aarhus and the EU Commission (EU-Biotech and TMR Programmes) (S.N.).

Correspondence and requests for materials should be addressed to P.A. (e-mail: pagre@jhmi.edu).

Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene

Guo-Liang Xu*, Timothy H. Bestor*, Déborah Bourc'h†, Chih-Lin Hsieh‡, Niels Tommerup§, Merete Bugge§, Maj Hulten||, Xiaoyan Qu†, James J. Russo† & Evani Viegas-Péquignot†

* Department of Genetics and Development, College of Physicians and Surgeons of Columbia University, New York 10032, USA

† INSERM U383, Hôpital Necker-Enfants Malades, 149 rue de Sèvres, 75743 Paris Cedex 15, France

‡ Department of Urology, and Department of Biochemistry and Molecular Biology, University of Southern California School of Medicine, 1441 Eastlake Avenue, Los Angeles, California 90033, USA

§ Department of Medical Genetics, Institute of Medical Biochemistry and Genetics, University of Copenhagen, Glostrup DK-2600, Denmark

|| Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, United Kingdom

¶ Columbia Genome Center, College of Physicians and Surgeons of Columbia University, New York 10032, USA

The recessive autosomal disorder known as ICF syndrome^{1–3} (for immunodeficiency, centromere instability and facial anomalies; Mendelian Inheritance in Man number 242860) is characterized by variable reductions in serum immunoglobulin levels which cause most ICF patients to succumb to infectious diseases before adulthood. Mild facial anomalies include hypertelorism, low-set ears, epicanthal folds and macroglossia. The cytogenetic abnormalities in lymphocytes are exuberant: juxtacentromeric heterochromatin is greatly elongated and thread-like in metaphase chromosomes, which is associated with the formation of complex multiradial chromosomes. The same juxtacentromeric regions are subject to persistent interphase self-associations and are extruded into nuclear blebs or micronuclei. Abnormalities are largely confined to tracts of classical satellites 2 and 3 at juxtacentromeric regions of chromosomes 1, 9 and 16. Classical satellite DNA is normally heavily methylated at cytosine residues, but in ICF syndrome it is almost completely unmethylated in all tissues⁴. ICF syndrome is the only genetic disorder known to involve constitutive abnormalities of genomic methylation patterns. Here we show that five unrelated ICF patients have mutations in both alleles of the gene that encodes DNA methyltransferase 3B (refs 5, 6). Cytosine methylation is essential for the organization and stabilization of a specific type of heterochromatin, and this methylation appears to be carried out by an enzyme specialized for the purpose.

Figure 1A shows the marked cytogenetic abnormalities that are characteristic of lymphocytes from ICF patients. Juxtacentromeric heterochromatin of chromosomes 1, 9 and 16 is greatly elongated and prone to breakage and rejoining, which leads to the formation of complex multiradial chromosomes (Fig. 1A, Ba). Breakage and

rejoining is mostly restricted to classical satellite DNA of lymphocytes⁷. There are no notable increases in the incidence of neoplasia or cutaneous abnormalities, as occurs in other chromosome breakage syndromes, although immunodeficiency of unknown origin is also characteristic of those disorders. Figure 1Bb shows extrusion of classical satellite DNA into nuclear blebs of ICF lymphocytes⁸. Condensations of classical satellite DNA in interphase nuclei of control subjects stain intensely with antibodies to 5-methylcytosine (m^5C), whereas nuclei of ICF cells show only uniform nucleoplasmic staining (Fig. 1Ca).

Further abnormalities of genomic methylation patterns in ICF syndrome are shown in Fig. 1Cc and 1D. Classical satellite DNA in adult somatic tissues is normally heavily methylated, and the juxtacentromeric regions of chromosomes 1, 9 and 16 are the most prominent features in spreads of metaphase chromosomes stained with antibodies to m^5C (ref. 9). Treatment of cultured cells with the demethylating drug 5-azacytidine induces elongation of the centromeric or juxtacentromeric regions of the same chromosomes that are affected in ICF syndrome⁴ and it was proposed that

ICF syndrome might involve constitutive demethylation of classical satellite DNA⁴. Indeed, these sequences were almost completely demethylated in DNA from all tissues of ICF patients (Fig. 1D, lanes 1–4), thereby identifying the first human genetic syndrome to involve defective genomic methylation patterns⁴. Demethylation is mostly confined to classical satellite DNA, and total content of m^5C is not measurably reduced (data not shown). Although classical satellite DNA is demethylated in all tissues of ICF patients, abundant chromosome abnormalities have to date been seen only in phytohaemagglutinin-stimulated T cells. The methylation level of retroposons (which contain >90% of the total m^5C ; ref. 10) is only slightly reduced in ICF DNA¹¹ (Fig. 1D, lanes 5–8), and allele-specific methylation is retained at the imprinted *H19* gene (Fig. 1D). Methylation of CpG islands on the inactive X chromosomes is reduced in ICF syndrome, although there are no notable defects in X-chromosome inactivation¹².

Transmission genetics¹³ and somatic cell hybridization experiments¹⁴ identified ICF syndrome as a recessive autosomal disorder. Homozygosity mapping (aided by the fact that the parents

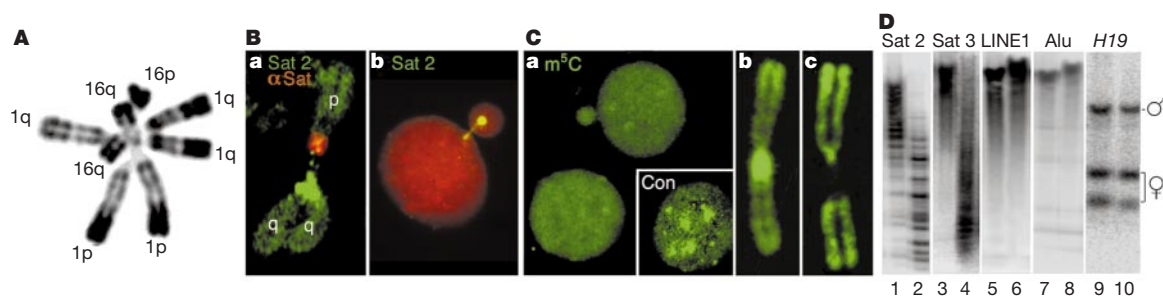


Figure 1 Cytogenetic and methylation abnormalities in ICF syndrome. **A**, Multiradiate chromosome from ICF lymphocyte with multiple p and q arms derived from chromosomes 1 and 16. Prepared for microscopy by R banding. **Ba**, Combined fluorescence *in situ* hybridization (FISH) for alphoid satellite (α Sat, red) and classical satellite 2 (Sat2, green) on a chromosome 1 that has undergone duplication of long (q) arms by breakage and rejoining. Note that breakage and rejoining has occurred within satellite 2; alphoid satellite at the centromere is not involved. **Bb**, FISH for satellite 2 shows extrusion into micronucleus or nuclear bleb; note apparent continuity of satellite 2 between bleb and nucleus. Propidium iodide counterstain in red. **Ca**, Interphase nuclei stained with antibody to m^5C (ref. 29). Concentrations of classical satellite show intense staining as foci in control nucleus (inset), whereas nuclei of ICF lymphocytes show uniform nucleoplasmic

staining. Upper nucleus shows nuclear bleb. **Cb**, Immunofluorescence micrograph of chromosome 1 from unaffected individual showing staining for m^5C in highly condensed juxtacentromeric satellite 2 DNA in proximal long arm. **Cc**, extended juxtacentromeric region of chromosome 1 from ICF T cell shows little or no staining for m^5C . **D**, Demethylation of satellites 2 and 3 but near-normal methylation of LINE1 and Alu retroposons and the imprinted *H19* gene in ICF syndrome. DNA from normal lymphoblastoid cells are in odd lanes, DNA from ICF is in even lanes. Maternal and paternal alleles of *H19* are in lanes 9 and 10. The methylation-sensitive restriction enzymes were *Bst*BI in lanes 1 and 2 and *Hpa*II in lanes 3–8; lanes 9 and 10 were double digests with *Hpa*II and *Rsa*I.

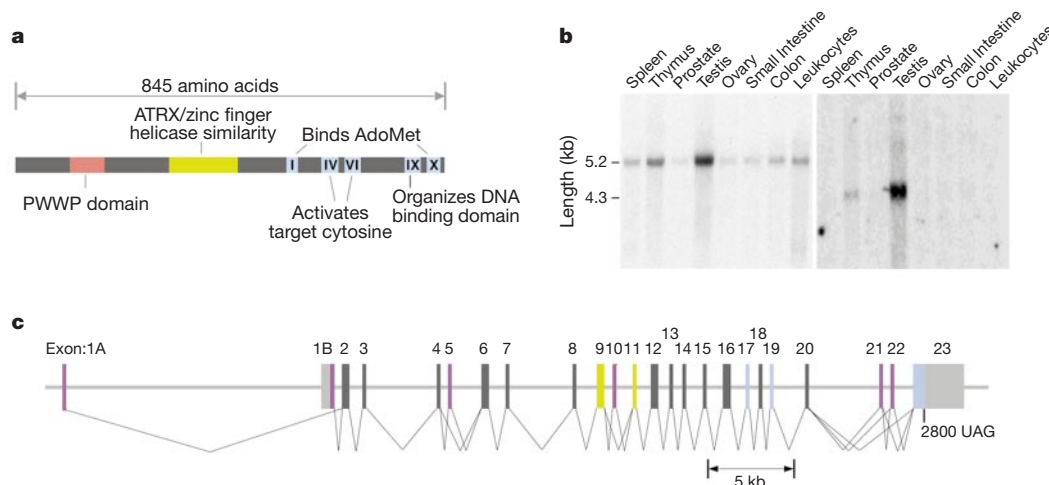


Figure 2 Structure and expression of *DNMT3B* gene and protein. **a**, Sequence motifs of *DNMT3B* protein. Motifs I, IV, VI, IX and X have known functions in the DNA methylation reaction¹⁹; PWWP and ATRX motifs in the N terminus are of undefined function. **b**, Analysis of *DNMT1* and *DNMT3B* mRNA expression in adult human tissues. The *DNMT3B* blot was exposed ten times longer than the *DNMT1* blot. Control probes

confirmed equal loading of all lanes (data not shown). **c**, Intron–exon structure of *DNMT3B* gene from 20q11. Exons that are normally alternatively spliced are shown in red, and exons that encode DNA methyltransferase and other motifs are coloured as in **a**. The major mRNA encodes an open reading frame of 845 amino acids, as shown in **a**.

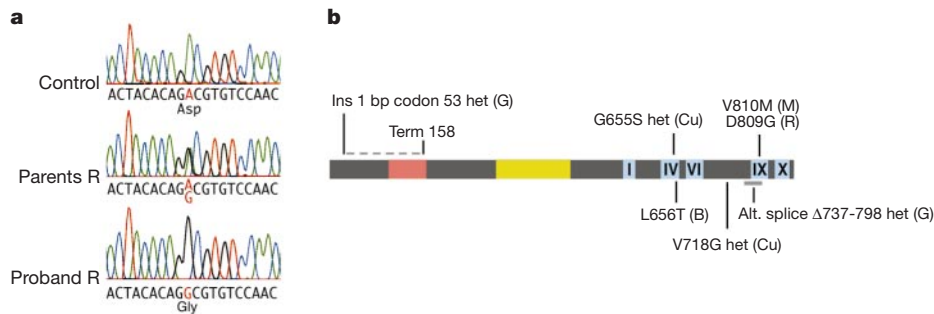


Figure 3 Detection of mutations in *DNMT3B* in ICF patients. **a**, D809G mutation in patient R. Both consanguineous parents were heterozygous for this mutation; analysis from father is shown. **b**, Summary of *DNMT3B* mutations in ICF syndrome. Family code appears in parentheses; heterozygous mutations are designated het. Term, termination. All

of two affected sibs were fifth cousins) was used to map the ICF trait to 20q11-13 (ref. 13). This region contains the *DNMT3B* gene (Unigene identification number 96323)¹⁵, which had been identified as an expressed sequence tag (EST) that contains sequence motifs diagnostic of DNA methyltransferases^{5,6}. *DNMT3B* is more closely related to the multispecific DNA methyltransferases of *Bacillus* bacteriophages than to other eukaryotic DNA methyltransferases. Low levels of enzymatic activity in purified Dnmt3b of mouse have been reported⁵ and *DNMT3B* is thought to be a *de novo* enzyme⁵, in adherence to the traditional view that *de novo* and maintenance methylation (the specific methylation of hemimethylated DNA during S phase) must be mediated by separate enzymes¹⁶.

We assembled the complete *DNMT3B* complementary DNA from overlapping partial clones extracted from a testis cDNA library. Characterization of the *DNMT3B* gene revealed that it spans ~47 kilobases (kb) and contains 24 exons, 6 of which are subject to alternative splicing (Fig. 2c). There are two alternative 5' exons. The human *DNMT3B* messenger RNA is 4.3 kb long and is mostly restricted to testis and thymus; levels of mRNA expression in adult tissues are far less than levels of *DNMT1* (Fig. 2b). The carboxy-terminal region contains five of the most highly conserved DNA methyltransferase motifs^{17,18}, whose functions in the trans-methylation reaction are known from crystallography data¹⁹. The amino-terminal region of *DNMT3B* contains a cysteine-rich domain that is distinct from the Zn-binding domain of *DNMT1* but is closely related to a region of the putative DNA helicase XH2 or ATRX, which is inactivated in the α -thalassemia-mental retardation syndrome known as ATRX²⁰. Closer to the N terminus of *DNMT3B* is a PWWP domain²¹ of unknown function (Fig. 2a).

When the ICF critical region was found to contain the gene for a DNA methyltransferase, studies were undertaken to determine whether mutant alleles of *DNMT3B* occur in ICF patients. The clinical diagnosis of ICF syndrome was confirmed in each of five unrelated patients by observation of the unusual cytogenetic abnormalities associated with variable combined immunodeficiency. DNA samples from patients were demethylated at classical satellites 2 (from chromosomes 1 and 16) and 3 (from chromosome 9; ref. 22). Families Cu and G were not known to be consanguineous and were heterozygous for markers between D20S850 and D20S200, which flank the ICF locus and the *DNMT3B* gene. Families M, B and R were known to be consanguineous and were found to be homozygous for markers in the ICF critical region. Despite a common origin in Jutland, patients M and B were found to have different haplotypes in this region (data not shown). Mutation detection involved polymerase chain reaction (PCR) amplification of each *DNMT3B* exon and direct sequence determination. Each patient was found to be a homozygote or compound heterozygote with respect to mutations that affected the conserved DNA methyltransferase sequence motifs. As shown in Fig. 3a, patient R is homozygous for an aspartic acid to glycine substitution within

mutations affect residues that are invariant in *DNMT3A* and *DNMT3B* of mouse and human, and in an uncategorized zebrafish Dnmt3 family member. Amino-acid number is with respect to the open reading frame in accession code AF176228.

motif IX, a highly conserved motif that is involved in providing structure to the variable sequence recognition domain of DNA methyltransferases¹⁹. The consanguineous parents were both heterozygous for this mutation. None of 25 normal individuals screened by direct sequencing showed sequence alterations in this region, nor did the mutation occur in a further 30 individuals screened for the *Afl*III restriction site that is removed by the mutation. Patients M, B, Cu and G also had mutations in DNA methyltransferase motifs (Fig. 3b), and, as predicted by haplotype analysis, patients M and B had different mutations. Patient Cu was a compound heterozygote for missense mutations, and patient G was a compound heterozygote in which the maternal allele of *DNMT3B* had an aberrant alternative splice that removed exons 21 and 22 and amino acids 737–798 (Fig. 3b), whereas the paternal allele had a

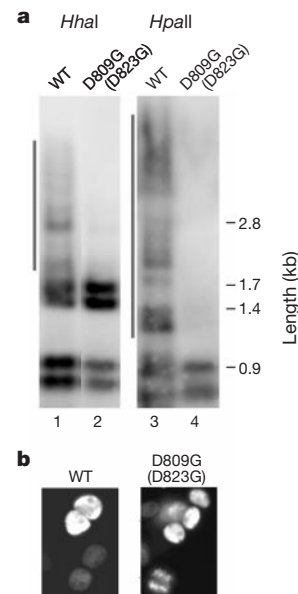


Figure 4 The D809G mutation found in family R eliminates enzymatic activity of *DNMT3B*. **a**, Propagation of an episome that encodes the wild type (WT, lanes 1 and 3) or D809G variant (lanes 2 and 4) of *DNMT3B* led to methylation of a co-transfected assay episome only in the case of the wild type, confirming that the mutation eliminates or greatly reduces enzymatic activity. Methylated fragments are marked with vertical bars to the left of each autoradiogram. The mouse cDNA (identical in this region to the human sequence) was used; the D809G mutation in the human sequence corresponds to the D823G mutation in the mouse sequence shown in parentheses. A mutation that removes the active-site cysteine residue¹⁹ also eliminates enzyme activity²³. **b**, Confirmation that wild-type and mutant proteins were expressed in nuclei at comparable levels in transfected cells. Cells were fixed and stained with monoclonal antibody to an N-terminal c-MYC tag as described²³.

single nucleotide insertion at codon 53 that caused a frameshift and early termination. The maternal allele had three intronic mutations that could have caused the alternative splicing (G65452A and T65903C in intron 21, and G67577A in intron 22; numbering from GenBank accession code AL035071, which covers the entire *DNMT3B* gene), although the cause of the aberrant splicing was not assigned to a particular mutation. In every case, the substitutions and alternative splicing affected positions that are invariant in mouse and human *DNMT3B*, in the paralogue *DNMT3A* on chromosome 2, and in a zebrafish member of the *DNMT3* family.

The ICF patients were homozygotes or compound heterozygotes for mutations that affect DNA methyltransferase motifs, and all parents were heterozygous for mutant and wild-type alleles (with the exception of patient Cu, for whom parental DNA could not be obtained). All mutations affected positions that are invariant in all members of the DNMT3 protein family, which indicates that defective DNA methyltransferase activity is responsible for the syndrome. This was confirmed when the D809G mutation of patient R greatly reduced DNA methyltransferase activity when introduced into the mouse *Dnmt3b* cDNA carried by an episomal expression construct (Fig. 4). Both mutant and wild-type *Dnmt3b* proteins were nuclear and were expressed at similar levels, as established by immunostaining of transfected cells that expressed *Dnmt3b* tagged with a c-MYC epitope at the C terminus (Fig. 4b). A mutation that converted the active-site cysteine into a serine eliminated all *de novo* methylation in this assay²³, confirming that the observed cytosine methylation is the direct result of *Dnmt3b* methyltransferase activity encoded by the expression construct and not due to alterations of chromatin structure or other effects²⁴. These data confirm that the ultimate cause of ICF syndrome is defective DNA methylation. It should be noted that none of the patients was homozygous for mutations that removed sequences N-terminal of the DNA methyltransferase motifs, which suggests an essential function for this region that may not be directly involved in the transmethylation reaction. The observed mutations do remove all detectable methylation from classical satellite DNA in ICF patients⁴, which indicates that they may be null alleles with respect to enzymatic activity.

The selective undermethylation of classical satellite DNA in ICF syndrome indicates that DNMT3B may be specialized for the methylation of these sequences, and the pronounced decondensation and abundant breakage and rejoining at classical satellite tracts of chromosomes 1, 9 and 16 in ICF lymphocytes suggest that cytosine methylation is required for their organization and stabilization. This in turn provides clues as to the source of immunodeficiency in ICF syndrome and the function of classical satellite DNA, which is retained in large amounts (>3% of the genome, or ~200 million bp per nucleus²²) despite the lack of a known function. Centromeric heterochromatin forms discrete foci in interphase nuclei that are enriched in heterochromatin proteins involved in transcriptional silencing, and silenced genes migrate into these foci²⁵. We propose that methylation may be required for the assembly of the silencing foci around condensation of satellite DNA, for the recruitment of silencing proteins to the foci, or for the migration of silent genes into the foci. The loss of methylation in ICF syndrome may interfere with one or more of these processes, and the resultant persistent expression of stage-specific genes could interfere with lymphocyte maturation and cause the severe immunodeficiency characteristic of this disorder. Global genome methylation is dependent on *Dnmt1*, and introduced reporter genes show elevated rates of deletions and rearrangements in *Dnmt1*-null embryonic stem cells²⁶. These and other data suggest that methylation-dependent sequestration events protect the genome against deletions, rearrangements and translocations mediated by homologous recombination between repeated sequences¹⁰. Cytosine methylation therefore appears to be important in genome stabilization and both *cis*²⁷ and *trans* promoter silencing in mammals. □

Methods

DNMT3B analysis and sequence determination

DNMT3B EST clones AA216697 and T66356 were identified by BLAST search with bacterial DNA methyltransferase *M.XorI* as query sequences. The EST clones were used to screen cDNA libraries prepared from human testis mRNA (Stratagene), and a full-length cDNA (accession code AF176228) was assembled from overlapping clones. The entire *DNMT3B* genomic coding region had been sequenced by the Sanger Center and deposited as accession code AL035071. Each of the *DNMT3B* exons shown in Fig. 2 was amplified by PCR with primers in flanking introns; sequences of the primers are available on request. PCR products were sequenced in both directions by the dye-terminator method on ABI 377 sequencers.

Oligonucleotide probes for satellites 2 and 3 in Fig. 1D and for satellite 2 in Fig. 1B were based on the reported consensus sequences²². The LINE1 probe corresponded to the first 300 nucleotides of L1.3 (accession code L19088). The *H19* probe covered the first 2 kb of coding sequence of the *H19* gene, and the Alu probe corresponded to the first 20 nucleotides of the consensus for recently transposed AluYa5 elements.

Episomal assay of DNMT3B enzyme activity

Details of the generation of the wild-type murine *Dnmt3B* episomal expression vector, pMT3bMyc, have been described²³. The D809G mutation (D823G in the murine *Dnmt3B* used) was introduced by site-directed mutagenesis. We confirmed the presence of the ICF mutation and the lack of additional mutations by sequencing both strands of the insert, and the expression of full-length proteins to similar levels and stability by immunostaining transiently transfected cells with monoclonal anti-MYC antibody 9E10. Equimolar amounts of the assay plasmid, p220.2 (ref. 23), and the pMT3bMyc or pICF3b expression plasmid were co-transfected into 293/EBNA1 cells, in which the plasmids replicate as episomes²⁸. After eight days of propagation, cells were harvested for episomal DNA extraction by the Hirt method²⁸. The assays were carried out without selection for the episomal plasmid. DNA from each transfected culture was digested with *HpaII* or *HhaI* restriction enzymes to determine methylation status. The digested DNA was fractionated on 0.8% agarose gels, transferred to nylon membranes and probed with the entire p220.2 plasmid.

Received 10 August; accepted 13 October 1999.

- Maraschio, P., Zuffardi, O., Dalla Fior, T. & Tiepolo, L. Immunodeficiency, centromeric heterochromatin instability of chromosomes 1, 9, and 16, and facial anomalies; the ICF syndrome. *J. Med. Genet.* **25**, 173–180 (1988).
- Hulten, M. Selective somatic pairing and fragility at 1q12 in a boy with common variable immunodeficiency. *Clin. Genet.* **14**, 294 (1978).
- Tiepolo, L. *et al.* Concurrent instability at specific sites of chromosomes 1, 9, and 16 resulting in multibranched chromosomes. *Clin. Genet.* **14**, 313–314 (1978).
- Jeanpierre, M. *et al.* An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome. *Hum. Mol. Genet.* **2**, 731–735 (1993).
- Okano, M., Xie, S. & Li, E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5)-methyltransferases. *Nature Genet.* **19**, 219–220 (1998).
- Robertson, K. D. *et al.* The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. *Nucleic Acids Res.* **27**, 2291–2298 (1999).
- Sumner, A. T., Mitchell, A. R. & Ellis, P. M. A FISH study of chromosome fusion in the ICF syndrome: involvement of paracentric heterochromatin but not of the centromeres themselves. *J. Med. Genet.* **35**, 833–835 (1998).
- Maraschio, P., Cortinovis, M., Dainotti, E., Tupler, R. & Tiepolo, L. Interphase cytogenetics of the ICF syndrome. *Ann. Hum. Genet.* **56**, 273–288 (1992).
- Lubit, B. W., Pham, T. D., Miller, O. J. & Erlanger, B. F. Localization of 5-methylcytosine in human metaphase chromosomes by immunoelectron microscopy. *Cell* **9**, 503–509 (1976).
- Yoder, J. A., Walsh, C. P. & Bestor, T. H. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13**, 335–340 (1997).
- Miniou, P. *et al.* Abnormal methylation pattern in constitutive and facultative (X inactive chromosome) heterochromatin of ICF patients. *Hum. Mol. Genet.* **3**, 2093–2102 (1994).
- Bourc'his, D. *et al.* Abnormal methylation does not prevent X inactivation in ICF patients. *Cytogenet. Cell Genet.* **84**, 245–252 (1999).
- Wijmenga, C. *et al.* Localization of the ICF syndrome to chromosome 20 by homozygosity mapping. *Am. J. Hum. Genet.* **63**, 803–809 (1998).
- Schuffenhauer, S. *et al.* DNA, FISH and complementation studies in ICF syndrome: DNA hypomethylation of repetitive and single copy loci and evidence for a trans acting factor. *Hum. Genet.* **96**, 562–571 (1995).
- Schuler, G. D. Pieces of the puzzle: expressed sequence tags and the catalog of human genes. *J. Mol. Med.* **75**, 694–698 (1997).
- Holliday, R. & Pugh, J. DNA modification mechanisms and gene activity during development. *Science* **187**, 226–232 (1975).
- Posfai, J., Bhagwat, A. S., Posfai, G. & Roberts, R. J. Predictive motifs derived from cytosine methyltransferases. *Nucleic Acids Res.* **17**, 2421–2435 (1989).
- Lauster, R., Trautner, T. A. & Noyer-Weidner, M. Cytosine-specific type II DNA methyltransferases. A conserved enzyme core with variable target-recognition domains. *J. Mol. Biol.* **206**, 305–312 (1989).
- Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, S. *HhaI* methyltransferase flips its target base out of the DNA helix. *Cell* **76**, 357–369 (1994).
- Gibbons, R. J., Picketts, D. J., Villard, L. & Higgs, D. R. Mutations in a putative global transcriptional regulator cause X-linked mental retardation with thalassemia (ATR-X syndrome). *Cell* **80**, 837–845 (1995).
- Stec, I. *et al.* WHSC1, a 90 kb SET domain-containing gene, expressed in early development and homologous to a *Drosophila* dysmorphia gene maps in the Wolf-Hirschhorn syndrome critical region and is fused to IgH in t(4;14) multiple myeloma. *Hum. Mol. Genet.* **7**, 1071–1082 (1998).
- Jeanpierre, M. Human satellites 2 and 3. *Ann. Genet.* **37**, 163–171 (1994).

23. Hsieh, C.-L. *In vivo* activity of murine *de novo* methyltransferases, Dnmt3a and Dnmt3b. *Mol. Cell. Biol.* (in the press).
24. Jeddlosh, J. A., Stokes, T. L. & Richards, E. Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nature Genet.* **22**, 94–97 (1999).
25. Brown, K. E. *et al.* Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **91**, 845–854 (1997).
26. Chen, R. Z., Pettersson, U., Beard, C., Jackson-Grusby, L. & Jaenisch, R. DNA hypomethylation leads to elevated mutation rates. *Nature* **395**, 89–93 (1998).
27. Bestor, T. H. The host defence function of genomic methylation patterns. *Novartis Found. Symp.* **214**, 187–195 (1998).
28. Hsieh, C.-L. Dependence of transcriptional repression on CpG methylation density. *Mol. Cell. Biol.* **14**, 5487–5494 (1994).
29. Reynaud, C. *et al.* Monitoring of urinary excretion of modified nucleosides in cancer patients using a set of six monoclonal antibodies. *Cancer Lett.* **61**, 255–262 (1992).

Acknowledgements

We thank members of the European ICF Consortium for providing patient materials; B. F. Erlanger for discussions; E. Li for providing cDNA clone pMT3B for mouse Dnmt3B; B. Tycko for DNA samples; L. Nickelsen for technical assistance; K. Anderson for comments on the manuscript; and A. Niveleau for monoclonal antibody to m⁵C. Supported by grants from the NIH and the Leukemia Society of America (T.H.B.) and the Danish Research Councils and the Danish Cancer Society (N.T.).

Correspondence and requests for materials should be addressed to T.H.B. (e-mail: THB12@columbia.edu).

A plant regulator controlling development of symbiotic root nodules

Leif Schauser, Andreas Roussis, Jiri Stiller* & Jens Stougaard

Laboratory of Gene Expression, Department of Molecular and Structural Biology, University of Aarhus, Gustav Wieds Vej 10, 8000 Aarhus C, Denmark

Symbiotic nitrogen-fixing root nodules on legumes are founded by root cortical cells that de-differentiate and restart cell division to establish nodule primordia. Bacterial microsymbionts invade these primordia through infection threads laid down by the plant and, after endocytosis, membrane-enclosed bacteroids occupy cells in the nitrogen-fixing tissue of functional nodules. The bacteria excrete lipochitin oligosaccharides^{1,2}, triggering a developmental process that is controlled by the plant and can be suppressed. Nodule inception initially relies on cell competence in a narrow infection zone located just behind the growing root tip. Older nodules then regulate the number of nodules on a root system by suppressing the development of nodule primordia³. To identify the regulatory components that act early in nodule induction, we characterized a transposon-tagged *Lotus japonicus* mutant, *nin* (for nodule inception), arrested at the stage of bacterial recognition. We show that *nin* is required for the formation of infection threads and the initiation of primordia. NIN protein has regional similarity to transcription factors, and the predicted DNA-binding/dimerization domain identifies and typifies a consensus motif conserved in plant proteins with a function in nitrogen-controlled development.

Inoculation of *L. japonicus*⁴ roots with *Mesorhizobium loti* bacteria results in the development of functional root nodules. The nitrogen reduced by the bacteroids in the infected nodule cells allows the plants to grow in the absence of a nitrogen source, such as nitrate (see refs 5, 6) (Fig. 1a). This provides conditions for a mutant screen in which mutants with ineffective symbiosis have a short

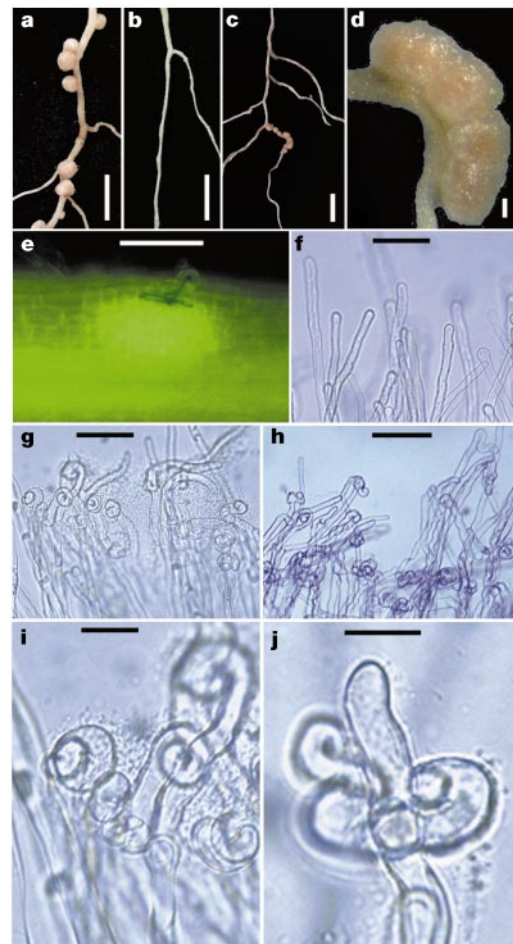
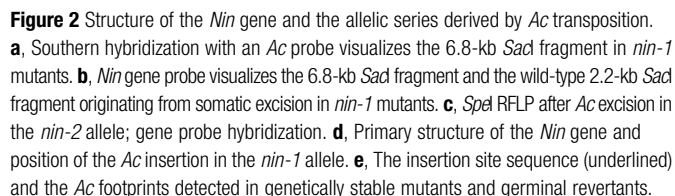


Figure 1 The phenotype of *nin* non-nodulating mutants. **a**, Wild-type root nodules. **b**, Non-nodulating *nin* phenotype. **c**, Nodulated revertant sector on an unstable *nin-1* mutant. **d**, Close up of a small revertant sector. **e**, Fluorescence microscopy of infection thread and nodule primordium on a wild-type plant. **f**, Light microscopy of root hairs on a *nin* mutant inoculated with the *nodC::Tn5* bacterial mutant. **g**, Root-hair curling in the wild type. **h**, Root-hair deformation and curling on a *nin* mutant. **i**, 'Shepherd's crook' in the wild type. **j**, Curling and deformation on a *nin* mutant. Scale bars: **a–c**, 1 cm; **d**, 1 mm; **e**, 0.9 mm; **f**, 50 μ m; **g**, 70 μ m; **h**, 100 μ m; **i**, **j**, 20 μ m.

stature, light-green chlorotic leaves and red stem pigmentation resulting from nitrogen starvation. Inoculated *nin* mutants did not develop nodules (Fig. 1b) and were therefore easily identified in a screen for transposon-tagged symbiotic mutants. After being transferred to nitrogen-containing nutrients, the *nin* mutants resumed growth. We did not observe any morphological change to the roots, shoots, leaves, flowers or seeds, indicating that *Nin* may not be required for general plant development.

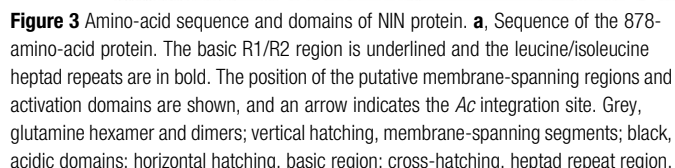
To establish when and where *Nin* acts during nodule development, bacterial invasion was visualized by staining for β -galactosidase activity expressed from a *hemA-lacZ* reporter in *M. loti*. An infection thread passing through a root hair towards the dividing cells of a wild-type plant is shown in Fig. 1e. Neither infection threads nor nodule primordia were observed in the *nin* mutants. Moreover, the first formative cell divisions in the outer cortex were not observed by light or fluorescence microscopy of *nin* mutants (not shown). After recognizing the microsymbiont, wild-type plants respond by curling root hairs to form the infection pockets of the infection zone (Fig. 1g, i). In *nin* mutants, normal curling into 'shepherd's crook' structures was accompanied by excessive deformation or curling of individual root hairs (Fig. 1h, j). In a plate inoculation system, the infection zone defined by root-hair curling covered ~2 mm in wild-type plants, compared with

* Present address: Plant Molecular Genetics, University of Tennessee, Knoxville, Tennessee 37901-1071, USA.



a

1 MEYGLLVQQQQQQDCNSAYGSLNSLSSDCGVTTHAEADHHIEELLVQGCWVEVGVGVREGELQLQDDESSFVGVKRRNWIGPAAAVAGSQSSQYKRL
101 VIAVGLYKDYTRNSNVLIQIWPVLRRLGLHDHYHTNLIYLSNNPPQPEAAADHESVSLGPFMPAAPNSLYSNVHVPFRFSHEYPRVQAQSGGSLALPV
201 FERGTGTCGLVLEIIVITQTITINYNVSNALDQAVDFRSSQSFIIPAIVKYVDELYQAAVNEIEVMTSVCKTHTNLPALTWAPCIQQKCGCGVSSSENYMW
301 CVSTVDSACFVGLDILGFGAEACEYHLFRQGGIVGTATFTSKPCFAIDITAFSAEYPLAHANMFLGHAAVAIPILRSVYTGSAADFVLEFFLPKDCDD
401 SEEGQQLNLSLMMVQQACRLSHVIVDEYITLPMPSHTSKELSEEERITITNNHEQKFLVSPSSHSEKSSSWIAHMEQAQKGGVSVSLYLELEP
501 KEFVKFTITNWDSSDTHDQQAQVFSDFSGQMSSEFGKASTVEGGDQIESSYTFGSRSSSSGGKRSSEKRRTKAEKTIISLOVLROYFAGSLDKAASKIGVCP
601 LKRICRQHGIITRPWSRKIKKVHSLKLQLVDSVQGAEGAIQIGSFYAPFELSSDFASACRSDDSSKKMHNYPDQNNLTLYGHGDHGVVTVSLKSPSPA
701 CSQTFAGNQPCFTIINNGDVLMTESPVPPEALLSRDRBCEAEALLNNASIQEDTKRFSRPKQETLPLSDSSGWNLSLETGAFRVKATFADEKIRFSLQPIW
801 GFSQDILAIARRFNLDVNTNILLKYLDLDDGEVWLACDGDGLECKDIRHSSQSRTIRLSLQFASPLNLNATFRNSSPS 878



Southern-blot analysis of mutant plants identified a common transposed *Ac* element, shown by the 6.8-kilobase (kb) *Sac*I fragment in Fig. 2a, b. In 50 normal and mutant T2 plants, this transposed *Ac* cosegregated with unstable mutant alleles present in heterozygotes and mutants, whereas it was absent in homozygous wild-type siblings and stable mutants. We used the transposon tag to isolate and subsequently sequence the *Nin* gene. In unstable mutants, a *Nin* gene probe (Fig. 2d) was used to detect the restriction fragment length polymorphism (RFLP) produced by an *Ac* (4.65 kb) insertion in the wild-type 2.2-kb *Sac*I fragment (Fig. 2b). This RFLP was absent in wild-type plants, in two stable mutants, and in 12 stable germinal revertants. The *Ac*-tagged

a

518 559 826

50 60 70 80 90

QGCWVEVSGVGVRREGELQLQDESSFFVVKRWIWGPAAAVAGSCNSSVKERL
EAAADHESVSLGPFMPAAPNSNLYSNVHVRFFRSHEYPRVQAQYQGSALP
VYDELYQAQAVNEIIEVMTSVCKTHNLPALTWAPCIQQQKCGCGVSSSEY
ITAFSAEYPLAHANMFMGLHAAVAIPLRSVYTGSAADFVLEFFLPKDC
TITNNHEQKLVFSPSSHESEKSLVAHMEAAQKQKGVSVLEYLEEP
TFGSRSSSSGGRRKSGEKRRTKAEKTIISLOVLROYFAGSLKDAAKSIGVCPTT
ASFPLESSSDFSASCRSDSSKKMHNYPDQNNLTLYGHGGGVVTSLKSPPSA
QEDTKFRSRPKSQTLPLPLSDSSGWNLSLETGAFRVKATFADEKIRFSLQPIW
RSSQSRTIRLSLFAQSPLNLANTFRNSSPS 878

R2

VCPTTLKRICROHGITRWPSRKIKKVVGHSLKKLQLVIDSVQCEGA
VCPTTLKRICROHGITRWPSRKIKKVVGHSLKKLQLVIDSVQCVQGS
VCPTTLKRICROHGIMRWPSRKIKKVVGHSLKKLQLVIDSVQCEGA
VCPTTLKRICROHGIMRWPSRKIKKVVGHSLKKLQLVIDSVQCEGA
VCPTTLKRICROHGIMRWPSRKIKKVVGHSLKKLQLVIDSVQCEGA
VCPTTLKRICROHGIMRWPSRKIKKVVGHSLKKLQLVIDSVQCEGA
VGLTVLKKKCRREFGPRWHRKKIKS DCLHDLQREAKKQCEKNEA

VSQTYLKKICRRLGDRWRFRRKIASVKKKH+
VSQTYLKKICRRLGDRWRFRRKIASVKKKH+

V--T-LK-ICR--GI-RWP-RKI--V--L
I

KR (12 x) RKxxKK
bipartite NLS

b, Alignment to the similar *Arabidopsis* proteins and the C terminus of the *Chlamydomonas* Mid proteins. Dots indicate heptad repeats; arrowhead indicates the K124 of Mid. Nuclear localization signals are shown. Helix predictions for NIN are shown as open boxes. Amino acids identical to NIN are shown on a black background and conservative changes on grey. The consensus includes amino acids, conserved in six or more proteins and I/L substitutions.

unstable mutant allele was designated *nin-1*. The structure of *Nin* and the position of *Ac* in exon 4 of the *nin-1* allele are shown in Fig. 2d.

In the *nin-1* allele, the *Ac* element was flanked by the characteristic 8-base pair (bp) direct repeat of the target insertion sequence (Fig. 2e). Excision of *Ac* creates small sequence alterations at the integration site. These 'footprints' were amplified using the polymerase chain reaction and were determined in 12 independent genetically stable revertants and two stable mutants. The stable mutant alleles *nin-2* and *nin-3* revealed sequence alterations causing translational frameshifts (Fig. 2e). In the *nin-2* allele, the footprint creates a *SpeI* restriction site and the RFLP seen in Fig. 2c. The footprints of the stable revertant alleles *Nin-4* and *Nin-5* translate into two and one additional amino acids, respectively (Fig. 2e). The wild-type sequence was found in 10 stable revertants. This allelic series with *Ac* in the unstable *nin-1* allele, footprints leading to stable frameshift mutations, and restoration of the reading frame in germinal revertants, demonstrates that the non-nodulating *nin* mutation was caused by insertion of the *Ac* element.

To further investigate the role of *Nin* in the molecular mechanisms governing nodule organogenesis, a complementary DNA was isolated and the structure of the derived protein analysed. The full-length *Nin* cDNA has an open reading frame of 2,634 bp, a 5' leader with stop codons in all three reading frames, and a 3' untranslated region of 141 bp. The protein, comprising 878 amino acids, is shown in Fig. 3a along with a diagram of several domains with similarity to transcription factors. The carboxy-terminal half of the protein carries an unusual putative DNA-binding domain with a basic region followed by two heptad leucine or isoleucine repeats (R1 and R2); an amphipathic leucine-zipper structure was predicted for R2 by helical-wheel analysis (Fig. 3). This arrangement is similar to the DNA-binding/dimerization domains of the bZIP⁸ and bHLH/Z⁹ transcription factors, but the R1/R2 region does not comply with a strict consensus of either domain. The distance between the basic region and the zipper, for example, is longer than would be expected from the bZIP consensus. The presence of a proline helix breaker (Pro 597) in R1 also distinguishes NIN, and alignment to prokaryotic helix-turn-helix DNA-binding domains¹⁰ suggests a GV turn motif preceding Pro 597 (not shown). No known DNA-binding/dimerization domains combine these components, and the structure of the domain has yet to be resolved by crystallography. Two acidic domains located between residues 428–503 and 826–845 may function as transcriptional activation domains. Supporting the role as a transcription factor, NIN has SV40-like and bipartite nuclear localization signals¹¹ (Fig. 3). Less clear is the role of a hexaglutamine stretch between residues 9–14 and five double and eight single glutamine residues within the first 420 amino acids.

Two membrane-spanning segments were predicted for NIN (Fig. 3a). One of these segments has similarity to previously assigned

membrane-spanning regions in the *Rickettsia prowazekii* ATP/ADP translocase¹² and in the photosystem I apoprotein A1 (ref. 13). The significance of the NIN segments is supported by the prediction of similarly positioned membrane helices in two *Arabidopsis* orthologues. These segments may alternatively constitute hydrophobic domains in the folded proteins.

The crucial role of *Nin* during nodulation may be taken to indicate that this gene was maintained only in legumes to serve this particular developmental process. This is not the case, however, as BLAST searches identified four similar *Arabidopsis* genes (F23E12.170, F6P23.15, T19F06.16 and T01024.24) of unknown function. In the translation products there is a remarkable conservation of the NIN putative DNA-binding domain (Fig. 3b), although other regions are less conserved. Further evidence of this region having an unusual structure comes from comparison of NIN and the four *Arabidopsis* proteins to the protein translated from the *Arabidopsis* F8D20.100 gene and the Minus dominance (Mid) proteins from *Chlamydomonas*¹⁴. This comparison defines a short consensus for these otherwise diverse proteins (Fig. 3b). Because the most striking feature is the invariant RWPxRK motif, we tentatively call this structure the RWP–RK domain. Its importance was demonstrated by a mutation changing the consensus Lys 124 to isoleucine¹⁴, which inactivated the Mid protein.

The regional similarity observed between NIN and the Mid proteins of *Chlamydomonas*^{14,15} (Fig. 3b) may have functional significance. In *Chlamydomonas*, Mid is encoded in the mating-type locus and determines the development of minus gametes during gametogenesis. The *mid* gene controls the expression of *minus*-specific genes and represses the expression of the opposing *plus*-specific genes. It functions as a developmental regulator gene and is both necessary and sufficient to determine the mating type of *Chlamydomonas* cells¹⁴. Gametogenesis in *C. reinhardtii* is induced by nitrogen limitation¹⁶, which is also a prerequisite not only for root nodule initiation but also to prevent the suppression of nodule

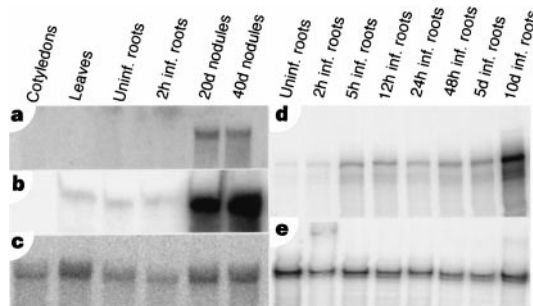


Figure 4 Expression of *Nin* during root-nodule development. **a, b**, *Nin* expression obtained by northern hybridization (**a**) and RNase protection studies (**b**) of different plant organs. **c**, Northern hybridization of Ubiquitin control. **d**, RNase protection showing *Nin* gene expression in root tissue after inoculation with *M. loti*. **e**, Ubiquitin control RNase protection. Time after infection (inf.) is shown in hours (h) or days (d).

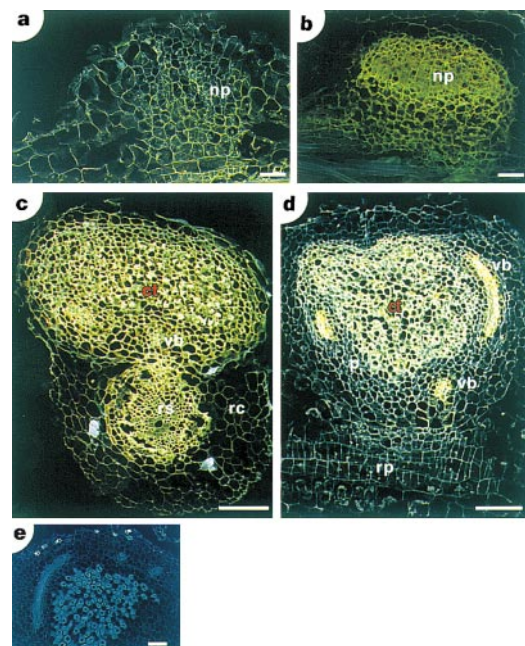


Figure 5 *In situ* localization of the *Nin* transcripts during nodule development. **a**, Early stage of module primordium with low-level *Nin* expression in dividing cells. **b**, Nodule primordium at a later stage expressing *Nin* at higher level. **c**, Transverse nodule section showing *Nin* expression in the central tissue, vascular bundles and root stele. **d**, Fully differentiated nitrogen-fixing nodule. *Nin* expression is confined to vascular bundles, parenchyma, and infected and uninfected cells of the central tissue. **e**, Sense control. Abbreviations: np, nodule primordium; ct, central tissue; rc, root cortex; rs, root stele; vb, vascular bundle. Scale bar, 100 μm.

development. The similarity between NIN and Mid is remarkable because the *Chlamydomonas mid* gene has undergone rapid evolution, and conservation was detected only between the closely related *C. reinhardtii* and *C. incerta*¹⁵.

The phenotype of the *nin* mutants indicates that *Nin* is involved in the inception of root nodules, and the features of NIN the protein suggest that it acts as a transcriptional regulator of genes required for nodule development. To determine the developmental stage and cell types in which *Nin* may be involved in the transcriptional regulation of downstream genes, we investigated *Nin* expression by northern hybridization, RNase protection and *in situ* localization. The highest number of transcripts was found in young pink nodules (20 days after infection) and mature nodules (40 days). Some steady-state messenger RNA was detected in cotyledons, leaves, stems and roots by the more sensitive RNase protection assay (Fig. 4a, b, c). The size of the *Nin* transcript was found to be approximately 3,000 nucleotides, corresponding to the 2,888-bp cDNA plus a poly(A) tail. To follow the kinetics of *Nin* expression, we determined the level of steady-state messenger RNA during early phases of nodule initiation, when root and nodule tissue cannot be analysed separately (Fig. 4d, e). A marginal upregulation was detected 5 h after infection, followed by a substantial increase 10 days after infection, when small nodules were clearly visible. *In situ* hybridization with antisense digoxigenin-labelled riboprobes was used to visualize the cellular expression pattern. In the developing nodules, transcripts were detected in the primordium cells (Fig. 5a, b, e). Later in nodule development, *Nin* expression was localized to the nodule parenchyma, the nodule vascular bundles, and the uninfected and infected cells of the central tissue (Fig. 5c, d), but was not detected or was significantly lower in the nodule and root cortex. In transverse sections, *Nin* transcripts were localized in root vascular tissue (Fig. 5c). This expression pattern shows that transcriptional regulation of *Nin* directs expression to certain cell types and that expression is not organ specific. Expression of *Nin* in mature nodules indicates that it may be involved not only during nodule inception, but also later during nodule development.

The lipochitin oligosaccharide-dependent root-hair curling and extensive root-hair deformation indicate that the perception of lipochitin oligosaccharide is at least partly functional in the *nin* mutants. The competence of the root hair cells in the infection zone and the recognition and early interaction with the microsymbiont therefore seem to be unaffected. These findings place *Nin* in the signal perception–transduction or gene-activation pathway downstream of the early signal exchange between the symbionts. Further experiments are required to establish the exact function, but the structure of NIN points towards a role in gene activation. The presence of putative membrane-spanning helices in NIN is reminiscent of the membrane-bound SREBP and Notch regulators. In humans, the SREBP proteins, which regulate the transcription of genes involved in cholesterol uptake and biosynthesis, are bound to the membrane of the nucleus and endoplasmic reticulum¹⁷. A two-step proteolytic cleavage involving an activation protein and a peptidase^{18,19} releases the bHLH/Z transcriptional activator under sterol depletion. Cell-fate determination in *Drosophila*, determined by the Notch transmembrane receptor, is another example of signalling mediated by proteolytic cleavage. Ligand binding at the extracellular domain induces the release of the intracellular domain, which enters the nucleus as a transcriptional co-regulator²⁰. Similarly, the membrane-spanning regions suggested for NIN raise the possibility of post-translational control, and a model for the role and regulation of *Nin* can be summarized as follows. Membrane-bound NIN in target cells is proteolytically cleaved, and relocalization of the C-terminal transcriptional regulator peptide to the cell nucleus allows a quick response to the signal transduction triggered by bacterial infection. Transcriptional activation of *Nin* takes place simultaneously in the dividing cells. Downstream genes are then activated and the nodule primordia are established. Another *Nin*-

dependent signal promotes the growth of infection threads and negatively regulates root-hair competence. Absence of this signal in mutants leads to the lack of infection thread, excessive root-hair response and expansion of the infection zone seen in *nin* plants.

The nitrogen regulation of nodule organogenesis and *Chlamydomonas* gametogenesis together with the conserved domain embedding the RWP–RK motif of NIN and Mid indicate that this domain is involved in the regulation of genes controlled by nitrogen status. Because the conserved domain was found only in proteins from algae and higher plants, this class of regulator proteins probably evolved after the common ancestor of algae and plants diverged from the other eukaryotes.

Methods

Plant material

The *Ac* transposon approach has been described²¹ and the screening conditions for nodulation with *M. loti* strain NZP2235 were as previously outlined²². The *nin* mutants were found in progeny from a T1 plant during a screen for symbiotic mutants segregating from 100 primary transformants carrying *Ac* launching constructs. The T2 segregation ratio was 644 wild type: 179 *nin* mutants. Homozygous and heterozygous plants confirming the recessive inheritance were identified after screening progeny from T2 and T3 plants. The *nin* mutant phenotype was tested and confirmed with another *M. loti* strain, R7A²³.

Microscopy

Plants were inoculated with *M. loti* NZP2235 carrying a *hema*–*lacZ* reporter construct. Entire roots were stained for *lacZ* activity²⁴, cleared in lactic acid and observed under an Axioscope microscope (Zeiss) with bright field or fluorescence. The *nodC::Tn5* strain is an R7A derivative.

Cosegregation analysis

DNA from mutant and normal siblings was digested and transferred to nylon membranes. The radioactive *Nin* gene probe was prepared by labelling of a PCR fragment using the Random Primed Kit (Boehringer Mannheim).

Isolation of flanking regions

The plant DNA flanking the left end of *Ac* in the *Nin-1* allele was isolated by inverse PCR²⁵ using the *Ac* primers 5'-CGACTTAAACCCGACCGATCCTATCGG-3' and 5'-GCAAGGCGCATGCATCAAAACAAGGCG-3'. Plant DNA flanking the right end was amplified using the *Ac* primer 5'-GTAGAGCTAGTTCCCGACCGTTTCACC-3' and a primer 5'-CCTCCATCATGTGTGCAATCCATG-3' derived from the cDNA. The DNA-extraction procedure has been described²². *Ac* excision footprints were obtained using the *Nin* gene-specific primers 5'-CAAGGAATTGTTGGTACAGCCTTC-3' and 5'-CCTCCATCATGTGTGCAATCCATG-3' flanking the *Ac* element. PCR fragments of 550 bp were sequenced.

cDNA isolation

The *Nin* gene probe was used to screen 700,000 plaque-forming units of an unamplified oligo dT-primed nodule cDNA library. One hybridizing clone of 2,541 bp was found. The full-length cDNA was assembled by adding 304 bp isolated from a random-primed nodule cDNA library. 5' Rapid amplification of cDNA ends was used to map the transcription start sites corresponding to transcripts of 2,863 and 2,907 nucleotides.

In situ hybridization and RNA analysis

Ribonuclease protection assays were according to Ambion RPAII(TM) and MAXIScript(TM). A *L. japonicus* 0.26-kb ubiquitin fragment was used for normalization. A 0.6-kb *Nin* 5' fragment was subcloned into pBluescript to provide templates for T3 and T7 polymerase transcription of sense and antisense DNA. Digoxigenin labelling of RNA probes, tissue preparation and *in situ* hybridization were as described²⁶. Images were processed by using Photoshop 4.0 for presentation.

Computer analysis of sequence data

Databases were searched with BLAST (<http://www.ncbi.nlm.nih.gov/>) and secondary protein structures were predicted by PredictProtein (<http://www.embl-heidelberg.de/>). Membrane-spanning regions were predicted by SBASE (<http://base.icgeb.trieste.it/>), TMPRED (<http://www.ch.embnet.org/>), TopPred II (<http://www.biokemi.su.se>) and PredictProtein.

Received 19 May; accepted 26 August 1999.

1. Truchet, G. *et al.* Sulphated lipo-oligosaccharide signals of *Rhizobium meliloti* elicit root nodule organogenesis in alfalfa. *Nature* **351**, 670–673 (1991).
2. Spaink, K. *et al.* A novel highly unsaturated fatty acid moiety of lipo-oligosaccharide signals determines host specificity of *Rhizobium*. *Nature* **354**, 125–130 (1991).
3. Caetano-Anolles, G. & Gresshoff, P. M. Plant genetic control of nodulation. *Annu. Rev. Microbiol.* **45**, 345–382 (1991).
4. Handberg, K. & Stougaard, J. *Lotus japonicus*, an autogamous, diploid legume species for classical and molecular genetics. *Plant J.* **2**, 487–496 (1992).

5. Spaink, H. P., Kondorosi, A. & Hooykaas, P. J. J. (eds) *The Rhizobiaceae* (Kluwer, Dordrecht, 1998).
6. Long, S. R. *Rhizobium* symbiosis: Nod factors in perspective. *Plant Cell* **8**, 1885–1898 (1996).
7. López-Lara, I. M. et al. Structural identification of the lipo-chitin oligosaccharide modulation signals of *Rhizobium loti*. *Mol. Microbiol.* **15**, 627–638 (1995).
8. Landschulz, W. H., Johnson, P. F. & McKnight, S. L. The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science* **240**, 1759–1764 (1988).
9. Ferré-D'Amaré, A. R., Prendergast, G. C., Ziff, E. B. & Burley, S. K. Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. *Nature* **363**, 38–45 (1993).
10. Puyet, A., Ibanez, A. M. & Espinosa, M. Characterisation of the *Streptococcus pneumoniae* maltosaccharide MalR, a member of the LacI-GalR family of repressors displaying distinctive genetic features. *J. Biol. Chem.* **268**, 25402–25408 (1993).
11. Silver, P. A. How proteins enter the nucleus. *Cell* **64**, 489–497 (1991).
12. Plano, G. V. & Winkler, H. H. Identification and initial topological analysis of the *Rickettsia prowazekii* ATP/ADP translocase. *J. Bacteriol.* **173**, 3389–3396 (1991).
13. Lehmbeck, J. et al. Sequence of two genes in pea chloroplast DNA coding for 84 and 82 kD polypeptides of the photosystem I complex. *Plant Mol. Biol.* **7**, 3–10 (1986).
14. Ferris, P. J. & Goodenough, U. W. Mating type in *Chlamydomonas* is specified by *mid*, the Minus-dominance gene. *Genetics* **146**, 859–869 (1997).
15. Ferris, P. J., Pavlovic, C., Fabry, S. & Goodenough, U. W. Rapid evolution of sex-related genes in *Chlamydomonas*. *Proc. Natl Acad. Sci. USA* **94**, 8634–8639 (1997).
16. Beck, C. F. & Haring, M. A. Gametic differentiation of *Chlamydomonas*. *Int. Rev. Cytol.* **168**, 259–302 (1996).
17. Yokoyama, C. et al. SREBP-1, a basic-helix-loop-helix-leucine zipper protein that controls transcription of the low density lipoprotein receptor gene. *Cell* **75**, 187–197 (1993).
18. Brown, M. S. & Goldstein, J. L. The SREBP pathway: Regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell* **89**, 331–340 (1997).
19. Sakai, J. et al. Molecular identification of the sterol-regulated luminal protease that cleaves SREBPs and controls lipid composition of animal cells. *Mol. Cell* **2**, 505–514 (1998).
20. Schroeter, E. H., Kisslinger, J. A. & Kopan, R. Notch-1 signalling requires ligand-induced proteolytic release of intracellular domain. *Nature* **393**, 382–386 (1998).
21. Thykjaer, T., Stiller, J., Handberg, K., Jones, J. & Stougaard, J. The maize transposable element *Ac* is mobile in the legume *Lotus japonicus*. *Plant Mol. Biol.* **27**, 981–993 (1995).
22. Schaefer, L. et al. Symbiotic mutants deficient in nodule establishment identified after T-DNA mutagenesis of *Lotus japonicus*. *Mol. Gen. Genet.* **259**, 414–423 (1998).
23. Sullivan, J. T. et al. Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. *Proc. Natl Acad. Sci. USA* **92**, 8985–8989 (1995).
24. Boivin, C. et al. *Rhizobium meliloti* genes encoding catabolism of trigonelline are induced under symbiotic conditions. *Plant Cell* **2**, 1157–1170 (1990).
25. Earp, D. J., Lowe, B. & Baker, B. Amplification of genomic sequences flanking transposable elements in host and heterologous plants: a tool for transposon tagging and genome characterization. *Nucleic Acids Res.* **18**, 3271–3279 (1990).
26. Papadopoulos, K., Roussis, A. & Katinakis, P. Phaseolus ENOD40 is involved in symbiotic and non-symbiotic organogenetic processes: expression during nodule and lateral root development. *Plant Mol. Biol.* **30**, 403–417 (1996).

Acknowledgements

We thank C. Ronson for supplying the *M. loti nodC* mutant; H. Spaink for the purified *M. loti* lipochitin oligosaccharide; and A. Nielsen and S. Rye for assistance. This research was supported by the Danish Biotechnology Programme and the SJVF Whole Plant Physiology Initiative. J. Stiller was supported by the Danish Science Research Councils postdoc programme, and A.R. by the EU-TMR programme.

Correspondence and requests for materials should be addressed to J.S. (e-mail: stougaard@mbio.aau.dk). The *Nin* gene sequence and cDNA sequence have been deposited in the EMBL database (accession nos AJ238956 and AJ239041, respectively).

CaMKII regulates the density of central glutamatergic synapses *in vivo*

Christopher Rongo & Joshua M. Kaplan

Department of Molecular and Cell Biology, LSA 361, University of California, Berkeley, California 94720-3200, USA

Synaptic connections undergo a dynamic process of stabilization or elimination during development, and this process is thought to be critical in memory and learning and in establishing the specificity of synaptic connections¹. The type II calcium- and calmodulin-dependent protein kinase (CaMKII) has been proposed to be pivotal in regulating synaptic strength^{2–4} and in maturation of synapses during development⁵. Here we describe how CaMKII regulates the formation of central glutamatergic synapses in *Caenorhabditis elegans*. During larval development, the density of ventral nerve cord synapses containing the GLR-1

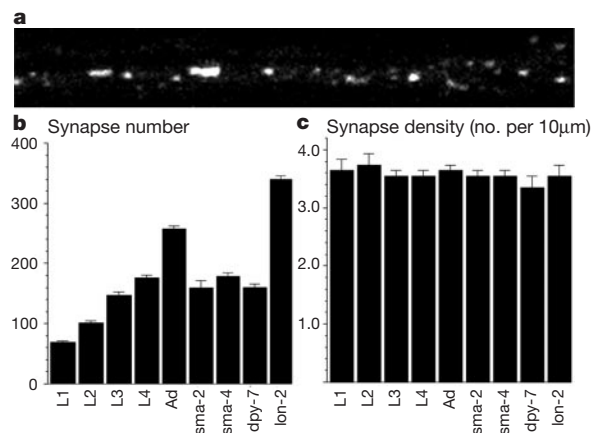


Figure 1 Formation of GLR-1::GFP clusters. **a**, GLR-1::GFP accumulates in clusters along the lengths of neurites in the ventral nerve cord. **b**, Individual GLR-1::GFP clusters in the ventral nerve cord were counted in wild-type larvae (L1–L4) and adults (Ad), as well as in mutant adults with short (*sma-2*, *sma-4* and *dpy-7*) or long (*lon-2*) body lengths. **c**, Data plotted as synaptic density (number of GLR-1::GFP clusters per 10 μm nerve cord length). Values are means ± s.e.m. Ten to thirty animals were examined for each entry.

glutamate receptor is held constant despite marked changes in neurite length. The coupling of synapse number to neurite length requires both CaMKII and voltage-gated calcium channels. CaMKII regulates GLR-1 by at least two distinct mechanisms: regulating transport of GLR-1 from cell bodies to neurites; and regulating the addition or maintenance of GLR-1 to postsynaptic elements.

To determine how central synapses change in number and morphology during development, we observed neuron–neuron synapses that contain the α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA)-type glutamate receptor (GluR) GLR-1 (refs 6, 7). We previously showed that chimaeric receptors tagged with the green fluorescent protein (GLR-1::GFP) can be used to visualize central glutamatergic synapses in living animals⁸. GLR-1::GFP is localized to discrete punctate structures (Fig. 1a; hereafter called clusters) along the lengths of neurites in both the ventral cord and nerve ring processes, where the majority of interneuron synapses reside⁹. GLR-1-containing neurites extend the full length of the ventral nerve cord throughout larval and adult development; thus, these neurites must elongate as the animal grows⁹. To see how synapse numbers are affected by changes in neurite length, we examined the number of GLR-1::GFP clusters in the ventral cord during larval development (Fig. 1b). First stage (L1) larvae had ~60 clusters whereas adults had ~260 clusters in the ventral cord. This apparent increase in GLR-1-containing synapses occurs despite the fact that the number of GLR-1-expressing cells remains constant throughout larval development^{6,7}. Homozygous *lin-5* mutants, which lack postembryonic cell divisions¹⁰, have similar numbers of GLR-1::GFP clusters (193 ± 17 , $n = 10$) as their age-matched (L4 stage) heterozygous siblings (181 ± 6 , $n = 10$). Therefore, the GLR-1::GFP clusters added during larval development are not induced by synaptic partners produced by post-embryonic lineages. One simple model explaining these results is that developing animals maintain a constant density of GLR-1-containing synapses along the ventral cord, thereby necessitating addition of synapses to compensate for increased body length.

We tested the relationship between synapse number and body length by comparing the number of GLR-1::GFP clusters with the length of the ventral cord in developing worms. Although cluster number varies nearly fourfold from early L1 to adulthood, the density of clusters was constant throughout larval development (~3.7 per 10 μm) (Fig. 1b, c). Similarly, compensatory changes occurred in the number of GLR-1::GFP clusters in short and long

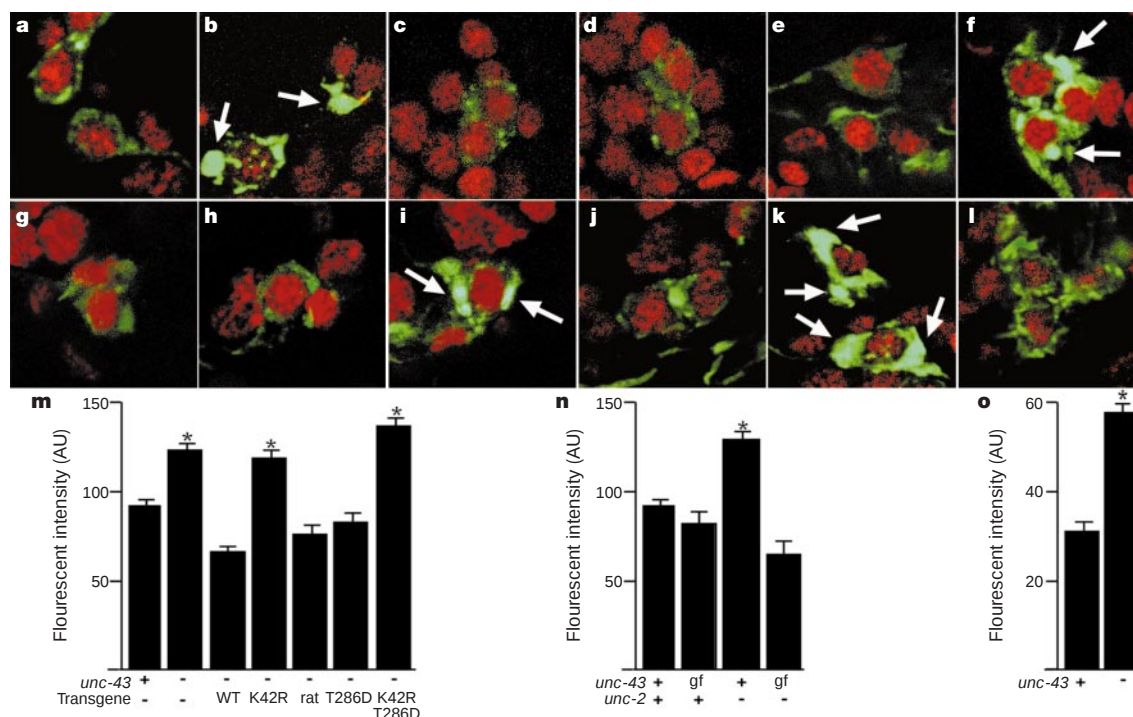


Figure 2 Mutants lacking the *unc-43* CaMKII accumulate GLR-1::GFP receptors in neuron cell bodies. Accumulation of GLR-1::GFP (**a, b, e–l**) or SNB-1::GFP (**c, d**), both shown in green, in the cell bodies of GLR-1-expressing interneurons was compared in various strains. Nuclei are labelled with propidium iodide (red). **a**, Wild-type. **b**, Homozygous *unc-43(lf)* CaMKII mutants carrying either no transgene, or *glr-1* promoted transgenes containing a wild-type UNC-43 cDNA (**e**), a kinase-defective mutant UNC-43(K42R) cDNA (**f**), a wild-type rat α CaMKII cDNA (**g**), a constitutively active UNC-43(T286D) cDNA (**h**), or a double mutant UNC-43(K42R, T286D) cDNA (**i**). **j**, Homozygous *unc-43(gf)* CaMKII

mutants. **k**, Homozygous *unc-2(lf)* calcium-channel mutants. **l**, *unc-43(gf)* CaMKII; *unc-2(lf)* double mutants. For comparison, equivalent amounts of SNB-1::GFP were found in the cell bodies of wild-type (**c**) and homozygous *unc-43(lf)* CaMKII mutant (**d**) interneurons. GLR-1::GFP (**m, n**) or endogenous GLR-1 (**o**) (detected by anti-GLR-1 antibodies²) cell-body fluorescence was quantified by measuring the pixel intensity of confocal projections. Values shown are means $\pm$ sem (AU, arbitrary units). From 19 to 55 nuclei were examined for each genotype. Asterisk, values significantly different ($P \leq 0.0001$) from wild type.

Table 1 UNC-43 regulates GLR-1 synaptic density

Effect of Ca ²⁺ and neural activity on GLR-1 synaptic density†		
Genotype	Gene product	Synapse density (no. per 10 μ m)
Wild type		3.7 $\pm$ 0.1
Synaptic transmission		
<i>unc-13</i>	Munc13	4.1 $\pm$ 0.1
<i>unc-18</i>	Sec1p	3.7 $\pm$ 0.1
<i>unc-64</i>	Syntaxin	3.8 $\pm$ 0.1
<i>unc-104</i>	SV Kinesin	4.0 $\pm$ 0.1
<i>snt-1</i>	Synaptotagmin	4.1 $\pm$ 0.1
Ca ²⁺ signalling		
<i>unc-43</i>	α CaMKII	2.3 $\pm$ 0.1*
<i>unc-2</i>	VDCC α 1	2.5 $\pm$ 0.1*
<i>unc-36</i>	VDCC α 2	2.8 $\pm$ 0.1*
<i>egl-19</i>	VDCC α 1	2.0 $\pm$ 0.1*
<i>egl-19(gf)</i>		2.7 $\pm$ 0.1*
<i>unc-2; egl-19</i>		2.3 $\pm$ 0.1*
<i>unc-2; unc-43</i>		2.2 $\pm$ 0.1*
UNC-43 regulates GLR-1 synaptic density		
Genotype	Transgene‡	Synapse density (no. per 10 μ m)
Wild type	–	3.7 $\pm$ 0.1
<i>unc-43</i>	–	2.4 $\pm$ 0.1*
<i>unc-43</i>	WT	3.6 $\pm$ 0.1
<i>unc-43</i>	K42R	2.5 $\pm$ 0.1*
<i>unc-43</i>	rat α	3.5 $\pm$ 0.1
<i>unc-43(gf)</i>	–	2.7 $\pm$ 0.1*
<i>unc-43</i>	T286D	3.1 $\pm$ 0.1*
<i>unc-43</i>	K42R, T286D	2.4 $\pm$ 0.1*
<i>unc-2</i>	–	2.5 $\pm$ 0.1*
<i>unc-2; unc-43(gf)</i>	–	2.1 $\pm$ 0.1*

Mutations are loss of function, unless gain of function is specified (gf). From 10 to 30 animals were examined for each genotype. Values are means $\pm$ s.e.m.

* Indicates a significant difference ($P < 0.0001$, t-test) from wild type.

† Synaptic density (number of GLR-1::GFP clusters per 10 μ m length of ventral nerve cord) was counted for each genotype.

‡ Transgene indicates the version of *unc-43* or rat CaMKII cDNA expressed by the *glr-1* promoter in the indicated genetic background. SV, synaptic vesicle; VDCC, voltage-dependent calcium channel; WT, wild type.

mutant adults, resulting in a constant density despite greatly differing adult body lengths (Fig. 1b, c). These results reveal a homeostatic mechanism coupling GLR-1::GFP synapse numbers and body length. A similar coupling of synaptic growth to increased size of the synaptic target is also seen at mouse and fly neuromuscular junctions¹¹. This phenomenon has been proposed as a mechanism to maintain relatively constant levels of postsynaptic excitation during developmental growth.

Are changes in synaptic transmission required to promote the increased number of GLR-1 synapses formed during development? We tested this idea by examining GLR-1::GFP clusters in mutants that have greatly reduced release of neurotransmitter, including *unc-13*, *unc-18* (nSec-1), *unc-64* (syntaxin), *snt-1* (synaptotagmin) and *unc-104* (kinesin)¹². None of these mutations reduced the number of GLR-1::GFP clusters in L1 larvae or young adults (Table 1). In fact, a modest increase in the density of GLR-1::GFP clusters was observed in some cases. These results are consistent with reports that a blockade of synaptic transmission often increases the number of postsynaptic elements in cultured rodent neurons^{13,14}.

CaMKII has been proposed to act as a critical signalling element regulating changes in synaptic strength in vertebrates^{2–4}. The *C. elegans* genome contains a single CaMKII gene, which corresponds to the *unc-43* gene¹⁵. Loss-of-function mutations in *unc-43* (*unc-43(lf)* alleles) reduce CaMKII activity, whereas a gain-of-function mutation (*unc-43(gf)*) causes constitutive calcium-independent activity¹⁵. Both *unc-43(lf)* and *unc-43(gf)* mutants had a significantly reduced density of clusters in adults (Table 1) but not in L1 larvae (data not shown). The cluster density in *unc-43(lf)* mutants was restored to wild-type levels by expressing a wild-

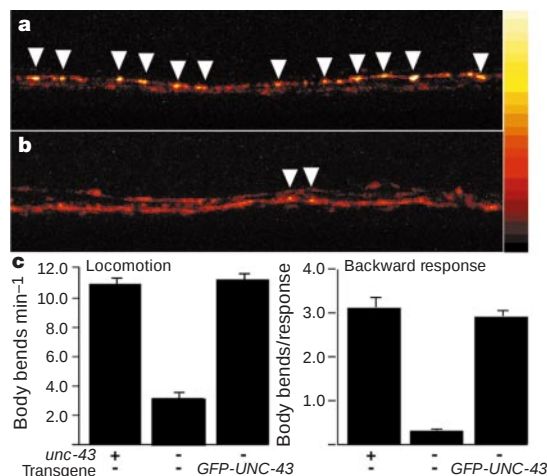


Figure 3 GFP::UNC-43 is localized to clusters in the ventral cord processes. **a**, **b**, Grey scale GFP::UNC-43 images were converted into colour according to the colour look-up table shown on the right. **a**, Wild-type neurons that express GFP::UNC-43 contain GFP::UNC-43 in clusters (arrowheads) in their neural processes. **b**, Wild-type neurons that express GFP::UNC-43 with a T286D substitution (GFP::UNC-43(T286D)) contain diffuse GFP::UNC-43(T286D) with few GFP::UNC-43(T286D) clusters (arrowheads) in their neural processes. **c**, Locomotion (worm body bends per min) and backward response to anterior body touch (worm body bends per touch) were counted for each genotype after a 1-h heat shock (33 °C). Transgene indicates whether *HS-GFP-unc-43* (GFP::UNC-43 expressed by the heat-shock promoter) was in the indicated genetic background.

type UNC-43 cDNA in the GLR-1-containing cells, whereas expression of a catalytically inactive mutant, UNC-43 (K42R), had no effect. Expression of a constitutively active, calcium-independent form of CaMKII, UNC-43(T286D)¹⁶, in the GLR-1-expressing interneurons also produced a significant reduction in the density of GLR-1::GFP clusters, whereas expression of the kinase-dead double mutant, UNC-43(K42R,T286D), did not (Table 1). These results indicate that CaMKII activity in GLR-1-expressing cells may regulate the addition of GLR-1-containing synapses in the ventral cord during larval growth.

What is the source of calcium that activates CaMKII? The *unc-2*, *unc-36* and *egl-19* genes encode subunits of voltage-dependent calcium channels^{17,18}. We found that loss-of-function mutations in *unc-2* and *egl-19* (and to a lesser extent *unc-36*) produced significant reductions in GLR-1::GFP clusters in adults (Table 1). Moreover, a gain-of-function mutation that is predicted to increase calcium influx, *egl-19(n2368sd)*¹⁸, also significantly reduced the density of GLR-1::GFP clusters in adults, similar to the defect seen in *unc-43(gf)* and *unc-43(T286D)* mutants. These results indicate that calcium influx through channels containing EGL-19 and UNC-2 subunits may regulate formation of GLR-1::GFP clusters in the ventral cord. Although both regulate cluster density, the calcium channels and *unc-43* CaMKII may regulate GLR-1 by different mechanisms. If this were the case, we would expect that double mutants lacking both CaMKII and the calcium channels would have more severe defects in cluster density than either single mutant. In contrast, we found that the cluster densities of the *unc-43(lf)*; *unc-2(lf)* double mutant and single mutants were indistinguishable (Table 1). These results indicate that the calcium channels and CaMKII form a single pathway regulating the density of GLR-1 synapses, as expected if calcium influx through these channels activates the endogenous *unc-43* CaMKII.

The effects of the *unc-43* and the calcium-channel mutations are both consistent with a model in which CaMKII plays both stimulatory and inhibitory roles in GLR-1 synapse formation. On the one hand, CaMKII activity promotes synapse formation, as seen by the decreased density of GLR-1::GFP clusters in mutants with either decreased calcium influx or decreased *unc-43* CaMKII activity. On

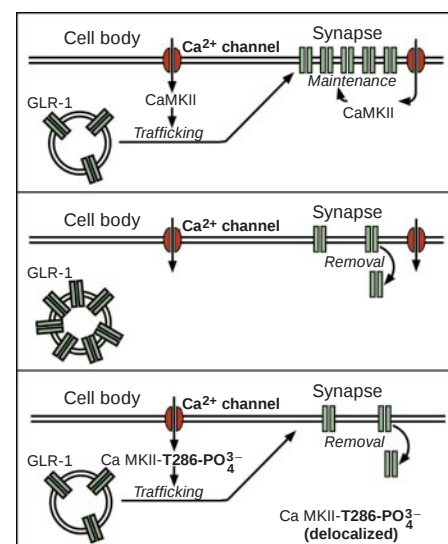


Figure 4 A model for CaMKII regulation of synaptic density. **a**, In wild-type worms, calcium influx through voltage-gated calcium channels and CaMKII activity promote transport of GLR-1 from the cell body to synapses, and addition or stability of GLR-1 at postsynaptic elements. **b**, In worms that lack *unc-43* CaMKII activity, GLR-1 accumulates in the cell body. **c**, Mutations that make *unc-43* CaMKII constitutively active result in proper trafficking of GLR-1, but fail to add or maintain GLR-1 at synapses, probably because the constitutively active form of CaMKII is not localized to synapses.

the other hand, CaMKII can also inhibit synapse formation, as seen by the decreased density of clusters in mutants with increased calcium influx or constitutive *unc-43* CaMKII activity. Although the underlying mechanisms might differ, there are precedents for both potentiation and depression of synaptic strength by CaMKII in vertebrates^{2,3}. CaMKII regulates growth of dendrites¹⁹; however, we found that the length of the interneuron neurites was not altered by changes in *unc-43* CaMKII activity (data not shown).

As a first step toward determining how *unc-43* CaMKII regulates the formation of GLR-1 synapses, we compared the subcellular localization patterns of GLR-1::GFP with a rescuing full-length GFP-tagged UNC-43 reporter (GFP::UNC-43)⁸. Both GFP-fusion proteins were found in ventral cord clusters, as well as in perinuclear structures in cell bodies (Figs 2a and 3a; and data not shown). Thus, *unc-43* CaMKII could regulate GLR-1::GFP in processes occurring either in the cell body (for example, intracellular trafficking) or in the axon (for example, addition of receptors to postsynaptic elements). The perinuclear GLR-1::GFP (Fig. 2a) is probably an intermediate in transport of GLR-1 to the axons, as transiently expressed GLR-1::GFP (utilizing a heat-shock vector) accumulated in this compartment at early times after heat-shock, but was found in axons at later time points (data not shown).

Mutants lacking either *unc-43* CaMKII or the calcium channels (*unc-2* and *egl-19*) accumulated high levels of GLR-1::GFP in the cell body, particularly in these perinuclear structures (Fig. 2b, k; and data not shown); however, the amount of GLR-1::GFP in individual ventral cord clusters of these mutants did not differ from that in wild-type animals (data not shown). Qualitatively similar results were observed when we compared the distribution of endogenously expressed GLR-1 (stained with anti-GLR-1 antibodies, Fig. 2o). Other synaptic proteins (for example, the vesicle protein SNB-1 (ref. 20)) did not accumulate in the cell bodies of *unc-43* mutants (82 ± 6 arbitrary fluorescence units (AU)), $n = 25$, compared with 98 ± 7 AU, $n = 33$ for wild type), indicating that CaMKII does not globally regulate the trafficking of neuronal membrane proteins (Fig. 2c, d). Expression of a wild-type UNC-43 complementary DNA in the GLR-1-expressing neurons rescued both the GLR-1 trafficking and the decreased synaptic-density defects, whereas

expression of the kinase-dead UNC-43(K42R) cDNA rescued neither (Table 1; Fig. 2e, f). Similarly, expression of a rat α CaMKII cDNA in the GLR-1-expressing neurons also rescued both the lowered synaptic density and the trafficking defects of *unc-43* mutants (Table 1; Fig. 2g), showing that this function of CaMKII has been conserved across phylogeny. Our results indicate that *unc-43* CaMKII regulates the trafficking of GLR-1 from the cell body to neurites and that CaMKII may possibly regulate the trafficking of AMPA receptors in mammalian neurons. These results are also consistent with reports that the delivery of AMPA receptors to postsynaptic elements in rodent neurons is regulated by neural activity^{13,14,21–23}.

Constitutive activation of *unc-43* CaMKII also resulted in reduced numbers of GLR-1-synapses; however, in this case, GLR-1::GFP did not accumulate in the cell bodies of neurons (Fig. 2h, j). The GLR-1 trafficking defect found in calcium-channel mutants was not observed in *unc-43(gf)*; *unc-2(lf)* double mutants (Fig. 2l), as expected if calcium influx through these channels activates *unc-43* CaMKII. These results also indicate that the normal calcium-dependent form of CaMKII activity (in wild-type animals) may have different physiological effects from the constitutively active version (which reduces synapse density). One model that might explain this is that these two forms of CaMKII are localized differently within neurons, and hence they have distinct effects. A GFP-tagged version of rat α CaMKII has been reported to be properly localized to synapses in cultured neurons²⁴, and we constructed a similar GFP-tagged form of UNC-43 to monitor its subcellular distribution *in vivo*. Expression of the GFP::UNC-43 construct corrected two behavioural defects observed in homozygous *unc-43(lf)* mutants (Fig. 3c), suggesting that fusion to the GFP moiety did not perturb the function or subcellular localization of *unc-43* CaMKII. Wild-type GFP::UNC-43 was localized in ventral cord clusters (Fig. 3a; 3.7 ± 0.4 clusters per $10 \mu\text{m}$, $n = 13$). In contrast, GFP::UNC-43(T286D) was diffusely distributed and had significantly ($P = 0.0001$) fewer ventral cord clusters (Fig. 3b; 1.3 ± 0.2 clusters per $10 \mu\text{m}$, $n = 11$). These two transgenes were expressed at similar levels (28 ± 3 AU, $n = 23$ for GFP::UNC-43, and 31 ± 3 AU, $n = 24$ for GFP::UNC-43(T286D)), indicating that the difference in localization pattern was not due to differences in expression levels. The *unc-43* CaMKII activity was not required for localization, as the kinase-dead GFP::UNC-43(K42R) protein was localized to clusters whereas a kinase-dead double mutant GFP::UNC-43(K42R, T286D) was not (data not shown). Thus, constitutively-activated CaMKII fails to localize to synaptic sites and also decreases the density of GLR-1::GFP clusters. These results indicate that synaptically localized calcium-dependent CaMKII activity may be required for either addition or maintenance of GLR-1::GFP at synaptic sites (Fig. 4a, b). Upon constitutive activation in response to high calcium levels, CaMKII is removed from synaptic sites and can no longer add or maintain GLR-1::GFP at these sites (Fig. 4c). These results are consistent with reports that synaptic localization of rat α CaMKII changes in response to autophosphorylation²⁴.

We have shown that CaMKII and voltage-dependent calcium channels are required for the homeostatic control of synaptic density during growth. We propose that this regulation is mediated by two distinct antagonistic functions of CaMKII (Fig. 4). As the worm grows in size, the synaptic density along the ventral cord drops, which should result in reduced excitation of the interneurons. Under these conditions, CaMKII induces formation of new synapses, perhaps by augmenting transport of nascent receptors from the cell bodies. Consistent with this idea, a slightly increased density of GLR-1::GFP clusters occurred in mutants that have reduced synaptic transmission, and hence should mimic reduced synaptic density (Table 1). Conversely, too high a synaptic density would cause increased excitation and constitutive CaMKII activity, which would destabilize existing synapses or prevent the

addition of new ones. One caveat for our results is that we have not directly measured changes in CaMKII activity in mutant or transgenic neurons. Thus, we cannot state unequivocally that the *unc-43(gf)* mutations or transgenes are actually changing the total kinase activity in these cells. Nonetheless, it is clear that *unc-43* CaMKII catalytic activity is required for maintaining the proper GLR-1 synaptic density and for GLR-1 trafficking. The mechanism by which CaMKII regulates GLR-1 trafficking and localization is not known but might involve direct phosphorylation of GLR-1^{25,26}.

Our results are similar in some respects to a form of plasticity, synaptic scaling^{11,27}. The synapses of cultured neurons are scaled up or down together, in response to changes in neural activity. Likewise, we have found that interneurons *in vivo* increase or decrease the number of postsynaptic elements, depending on their pattern of activity. Our results indicate that phenomena similar to synaptic scaling may occur *in vivo* and may be utilized to maintain constant density of excitatory synapses in response to dramatic growth of axons during development. □

Methods

Strains

dpy-7(e88), *egl-19(n582,ad695sd,n2368sd)*, *lin-5(e1348)*, *lon-2(e678)*, *sma-2(e502)*, *sma-4(e729)*, *snt-1(n2665)*, *unc-2(e55,e97,e129)*, *unc-13(e51)*, *unc-18(e81)*, *unc-36(e251)*, *unc-43(n1186,e408,n498sd)*, *unc-64(e246)* and *unc-104(e1265,rh142)*.

Molecular biology

The rat α CaMKII transgene (KP#350) was generated by ligating the rat α CaMKII cDNA (M. Kennedy) into the pV6 vector (V. Maricq), which contains the *glr-1* promoter and the *unc-54* 3'UTR. The *unc-43(T286D)* (KP#285), *unc-43(K42R)* (KP#351) and *unc-43(K42R, T286D)* (KP#352) transgenes were generated by modifying the *unc-43* cDNA by PCR, and verified by sequencing. The *GFP-unc-43(T286D)* (KP#287) transgene was generated by ligating the *unc-43(T286D)* cDNA into KP#241, a vector containing the *glr-1* promoter, a synthetic intron, GFP, and the *unc-54* 3'UTR. The *HS-GFP-unc-43* (KP#353) transgene was generated by ligating *GFP-unc-43* into PD49.78 (A. Fire).

Expression of transgenes

Transgenic strains were isolated by micro-injecting plasmids (typically at $50 \text{ ng } \mu\text{l}^{-1}$) using either *osm-10::gfp* (A. Hart) or *rol-6dm* (C. Mello) as a marker. KP#286 and KP#287 were injected at $20 \text{ ng } \mu\text{l}^{-1}$ without marker, and lines were identified by the transgene GFP fluorescence. Three or more lines were analysed in all transgenic experiments.

Analysis of fluorescence

GLR-1::GFP was observed in living animals as described⁸. For confocal microscopy, animals were fixed in 4% paraformaldehyde, then observed on a BioRad MRC1045. All animals were confocally sectioned in successive $0.5 \mu\text{m}$ focal planes under identical conditions to allow comparison with NIH Image. Image data filled the 8-bit dynamic range without saturation.

Received 28 June; accepted 3 September 1999.

- Hebb, D. *The Organization of Behavior* (Wiley, New York, 1949).
- Stevens, C., Tonegawa, S. & Wang, Y. The role of calcium-calmodulin kinase II in three forms of synaptic plasticity. *Curr. Biol.* **4**, 687–693 (1994).
- Mayford, M., Wang, J., Kandel, E. & O'Dell, T. CaMKII regulates the frequency-response function of hippocampal synapses for the production of both LTD and LTP. *Cell* **81**, 891–904 (1995).
- Lisman, J., Malenka, R. C., Nicoll, R. A. & Malinow, R. Learning mechanisms: the case for CaM-KII. *Science* **276**, 2001–2002 (1997).
- Wu, G., Malinow, R. & Cline, H. Maturation of a central glutamatergic synapse. *Science* **274**, 972–976 (1996).
- Hart, A., Sims, S. & Kaplan, J. A synaptic code for sensory modalities revealed by analysis of the *C. elegans* GLR-1 glutamate receptor. *Nature* **378**, 82–85 (1995).
- Maricq, A. V., Peckol, E., Driscoll, M. & Bargmann, C. *glr-1*, a *C. elegans* glutamate receptor that mediates mechanosensory signalling. *Nature* **278**, 78–81 (1995).
- Rongo, C., Whitfield, C. W., Rodal, A., Kim, S. K. & Kaplan, J. M. LIN-10 is a shared component of the polarized protein localization pathways in neurons and epithelia. *Cell* **94**, 751–759 (1998).
- White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of *Caenorhabditis elegans*. *Phil. Trans. R. Soc. Lond.* **314**, 1–340 (1986).
- Sulston, J. E. & Horvitz, H. R. Abnormal cell lineages in mutants of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **82**, 41–55 (1981).
- Turrigiano, G. G. Homeostatic plasticity in neuronal networks: the more things change, the more they stay the same. *Trends Neurosci.* **22**, 221–227 (1999).
- Riddle, D., Blumenthal, T., Meyer, B. & Priess, J. (eds) *C. elegans II* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1997).
- Rao, A. & Craig, A. M. Activity regulates the synaptic localization of the NMDA receptor in hippocampal neurons. *Neuron* **19**, 801–812 (1997).
- O'Brien, R. J. et al. Activity-dependent modulation of synaptic AMPA receptor accumulation. *Neuron* **21**, 1067–1078 (1998).

15. Reiner, D. J., Newton, E. M., Tian, H. & Thomas, J. H. Diverse behavioural defects caused by mutations in *Caenorhabditis elegans unc-43* CaM Kinase II. *Nature* **402**, 199–203 (1999).
16. Fong, Y. L., Taylor, W. L., Means, A. R. & Soderling, T. R. Studies of the regulatory mechanism of Ca²⁺/calmodulin-dependent protein kinase II. Mutation of threonine 286 to alanine and aspartate. *J. Biol. Chem.* **264**, 16759–16763 (1989).
17. Schafer, W. & Kenyon, C. A calcium-channel homologue required for adaptation to dopamine and serotonin in *C. elegans*. *Nature* **375**, 73–78 (1995).
18. Lee, R., Lobel, L., Hengartner, M., Horvitz, H. & Avery, L. Mutations in the $\alpha 1$ subunit of an L-type voltage-activated Ca²⁺ channel cause myotonia in *Caenorhabditis elegans*. *EMBO J.* **16**, 6066–6076 (1997).
19. Wu, G. Y. & Cline, H. T. Stabilization of dendritic arbor structure *in vivo* by CaMKII. *Science* **279**, 222–226 (1998).
20. Nonet, M., Saifec, O., Zhao, H., Rand, J. & Wei, L. Synaptic transmission deficits in *Caenorhabditis elegans* synaptobrevin mutants. *J. Neurosci.* **18**, 70–80 (1998).
21. Carroll, R. C., Lissin, D. V., von Zastrow, M., Nicoll, R. A. & Malenka, R. C. Rapid redistribution of glutamate receptors contributes to long-term depression in hippocampal cultures. *Nature Neurosci.* **2**, 454–460 (1999).
22. Liao, D., Zhang, X., O'Brien, R., Ehlers, M. D. & Huganir, R. L. Regulation of morphological postsynaptic silent synapses in developing hippocampal neurons. *Nature Neurosci.* **2**, 37–43 (1999).
23. Shi, S. H. et al. Rapid spine delivery and redistribution of AMPA receptors after synaptic NMDA receptor activation. *Science* **284**, 1811–1816 (1999).
24. Shen, K. & Meyer, T. Dynamic control of CaMKII translocation and localization in hippocampal neurons by NMDA receptor stimulation. *Science* **284**, 162–166 (1999).
25. Barria, A., Derkach, V. & Soderling, T. Identification of the Ca²⁺/calmodulin-dependent protein kinase II regulatory phosphorylation site in the α -amino-3-hydroxy-5-methyl-4-isoxazole-propionate-type glutamate receptor. *J. Biol. Chem.* **272**, 32727–32730 (1997).
26. Mammen, A. L., Kameyama, K., Roche, K. W. & Huganir, R. L. Phosphorylation of the α -amino-3-hydroxy-5-methylisoxazole-4-propionic acid receptor GluR1 subunit by calcium/calmodulin-dependent kinase II. *J. Biol. Chem.* **272**, 32528–32533 (1997).
27. Kim, J. H. & Huganir, R. L. Organization and regulation of proteins at synapses. *Curr. Opin. Cell Biol.* **11**, 248–254 (1999).

Acknowledgements

We thank D. Reiner and J. Thomas for sharing unpublished data. In addition, we thank M. Kennedy, H. Schulman, A. Fire, R. Barstead, A. Hart, J. Mendel, C. Mello and V. Maricq for providing plasmids and antibodies, and the *C. elegans* genome sequencing consortium for providing sequence. Some strains were provided by the *C. elegans* Genetics Center. We thank T. Serafini, E. Isacoff, G. Garriga and members of the Kaplan lab for comments on the manuscript. This work was supported by grants from the NIH and the Pew Charitable Trust (J.K.) and by a fellowship from Jane Coffin Childs (C.R.).

Correspondence and requests for materials should be addressed to J.K. (e-mail: joshkap@socrates.berkeley.edu).

Diverse behavioural defects caused by mutations in *Caenorhabditis elegans unc-43* CaM Kinase II

David J. Reiner*, Elizabeth M. Newton, Hong Tian & James H. Thomas

Department of Genetics, Box 357360, University of Washington, Seattle, Washington 98195, USA

Calcium/calmodulin-dependent serine/threonine kinase type II (CaMKII) is one of the most abundant proteins in the mammalian brain, where it is thought to regulate synaptic plasticity and other processes^{1–3}. Activation of the multisubunit kinase⁴ by calcium is effectively cooperative and can persist long after transient calcium rises^{5,6}. Despite extensive biochemical characterization of CaMKII and identification of numerous *in vitro* kinase targets¹, little is known about its function *in vivo*. Here we report that *unc-43* encodes the only *Caenorhabditis elegans* CaMKII. A gain-of-function *unc-43* mutation reduces locomotory activity, alters excitation of three muscle types and lengthens the period of the motor output of a behavioural clock. Null *unc-43* mutations cause phenotypes generally opposite to those of the gain-of-function mutation. Mutations in the *unc-103* potassium channel gene suppress a gain-of-function phenotype of *unc-43* in one tissue

without affecting other tissues; thus, UNC-103 may be a tissue-specific target of CaMKII *in vivo*.

The kinase activity of each subunit of the multisubunit CaMKII holoenzyme is activated by binding Ca²⁺-calmodulin. Activated subunits can phosphorylate adjacent activated subunits, causing a large increase in affinity for Ca²⁺-calmodulin and partial Ca²⁺ independence, properties that confer cooperativity and memory of Ca²⁺ spikes^{5,6}. CaMKII can phosphorylate a wide range of proteins *in vitro*, but under these conditions the kinase is nonspecific¹ and the functions of these potential targets *in vivo* are unclear. *In vivo* evidence for function has come from deletion of the CaMKII α gene in mouse, which compromises spatial learning in the hippocampus and reduces its synaptic correlate, long-term potentiation^{2,3}. Further progress has been impeded by the extreme complexity of the mammalian brain and the existence of four CaMKII genes in mammals with overlapping expression patterns⁷, making genetic redundancy likely.

unc-43 is defined by one gain-of-function (*gf*) mutation and many loss-of-function (*lf*) mutations^{8,9}. We genetically mapped *unc-43* to an interval that contains the single *C. elegans* CaMKII gene (Fig. 1). We sequenced the CaMKII coding region¹⁰ in twelve *unc-43* alleles and found mutations in each (Fig. 1c). Two of the *lf*

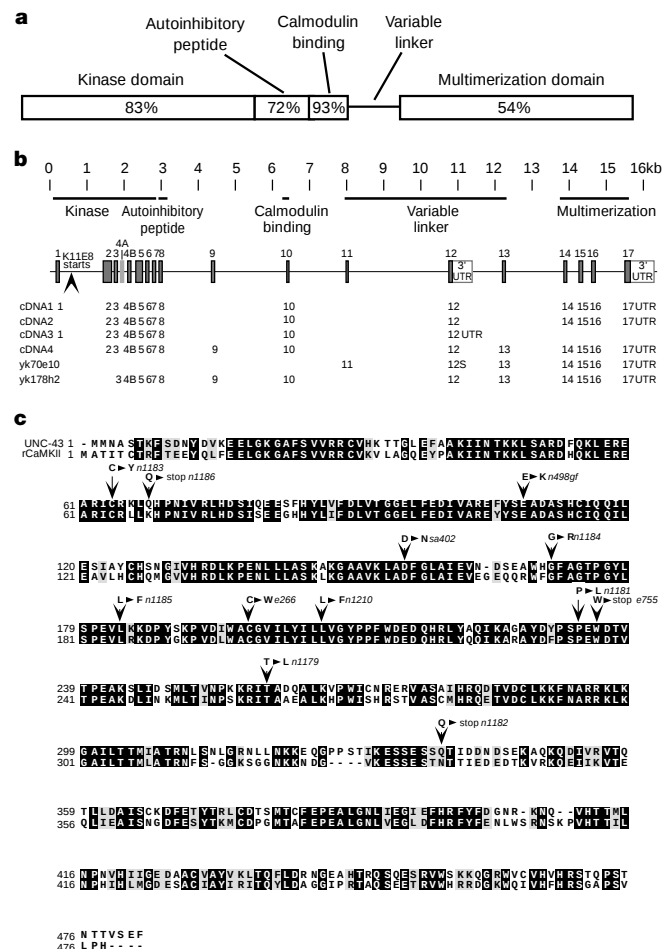


Figure 1 *unc-43* encodes CaMKII. **a**, cDNA1 has 69% overall identity with rat CaMKIIα. *unc-43* functional domains are shown in boxes with their percentage identity. **b**, *unc-43* alternate splicing. The yk70e10 exon 12, '12S', is six nucleotides shorter than exon 12 in other cDNAs. Exon 4A was not identified in a cDNA, but might encode a peptide very similar to the exon 4B peptide. A similar arrangement of alternative exons 4 is found in the *Caenorhabditis briggsae* CaMKII homologue (not shown)¹⁰. All *unc-43* exons, but not introns, share a high degree of identity with the *C. briggsae* homologue (not shown). **c**, Alignment of *C. elegans unc-43* cDNA1 with rat CaMKIIα. Black highlights identical residues, and grey highlights similar residues. Mutational changes are indicated.

* Present address: Howard Hughes Medical Institute and Department of Molecular and Cell Biology, University of California, Berkeley, California 94720-3204, USA

alleles, *n1186* and *e755*, result in early stop codons and are putative null mutations. The *n498gf* allele causes an E108K mutation in the kinase domain. The analogous E-to-R mutation in rat CaMKII α results in a kinase that is partially Ca^{2+} -independent (E. Yang and H. Schulman, personal communication), probably by destabilizing binding of the autoinhibitory domain. The structures of six *unc-43* transcripts were determined from complementary DNAs, revealing several splice variants between the calmodulin regulatory and multimerization domains (Fig. 1b). Similar variants have been found for *Drosophila* and mammalian CaMKII, in which they are proposed to influence subcellular localization, response to Ca^{2+} spike frequency and substrate specificity^{5,11,12,13}. One *unc-43* cDNA encodes a shorter CaMKII which lacks the multimerization domain, indicating that an unusual monomeric form of CaMKII may exist.

The UNC-43 expression pattern was determined by *in situ* antibody staining using 4A11, a monoclonal antibody to CaMKII from rat brain¹⁴. The antibody specifically stained the synapse-rich nerve ring, ventral nerve cord and dorsal nerve cord in wild-type but not in *unc-43(n1186)* or *unc-43(e755)* animals, indicating that the antibody specifically recognizes UNC-43 (Fig. 2). This observation and the absence of other CaMKII genes in the virtually complete genome sequence¹⁰ indicate that *unc-43* encodes the only CaMKII in *C. elegans*. Antibody staining did not distinguish neuronal from muscle expression at the neuromuscular junction because of their close apposition¹⁵, but a transgenic green fluorescent protein (GFP) fusion to the *unc-43* promoter was expressed widely in muscles and neurons and in the intestine (unpublished data). This broad expression pattern parallels that of CaMKII in mammals¹.

Mutations in *unc-43* cause multiple behavioural defects, includ-

ing complex perturbations in the behavioural clock controlling defecation. Defecation in *C. elegans* occurs by activation of a specific motor program every 45–50 seconds (Fig. 3a, c). This periodicity has properties that are characteristic of a clock, including partial temperature compensation¹⁶. The defecation motor program consists of two specific body-wall muscle contractions and an enteric muscle contraction (EMC), which expels gut contents^{17–19}. In the *unc-43(gf)* mutant, about 85% of defecation motor programs lacked the EMC step (see below), and the mean cycle period was longer and oscillated markedly over time (Fig. 3c, d, e). The *unc-43(gf)* cycle oscillation and EMC were correlated. Immediately after an EMC, the cycle was very long (88.2 ± 34.7 s), but as the animal became more constipated because of its EMC defect, the cycle period shortened markedly, finally becoming shorter than the wild type before another successful EMC (Fig. 3e). In contrast, *unc-43(lf)* caused an increased frequency of defecation, typified by a weaker repetition of the defecation motor program (that is, an echo) 10 s after the primary motor program (Fig. 3a, b), followed rarely by another echo 10 s after the first. As the latency of this echo motor program was highly reproducible (10.1 ± 1.4 s, $n = 70$), we propose that CaMKII suppresses a timed reactivation of the motor program. Defecation is initiated by an intestinal Ca^{2+} spike mediated by the inositol-trisphosphate-activated Ca^{2+} channel encoded by *itr-1* (*lef-1/dec-4*)²⁰. CaMKII may negatively regulate rhythmic ITR-1-mediated Ca^{2+} spikes, lengthening the cycle period in the *unc-43(gf)* mutant and permitting ectopic cycles in *unc-43(lf)*.

In addition to affecting the defecation clock, *unc-43(gf)* also caused defective EMCs (Fig. 4a), presumably by affecting either the

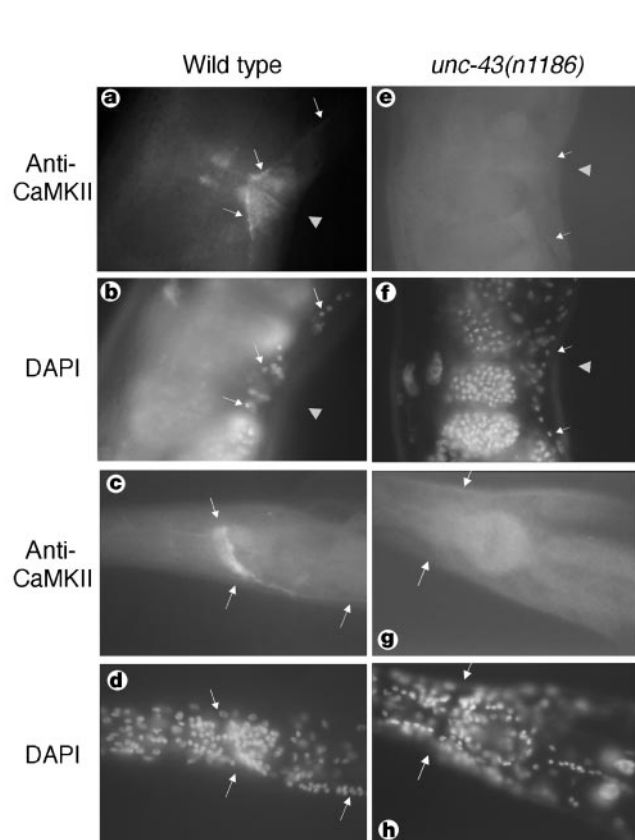


Figure 2 UNC-43 is expressed in the nervous system. Anti-CaMKII staining (a, c, e, g) of each animal is shown with DAPI staining beneath (b, d, f, h). a–d, Wild type; e–h, *unc-43(n1186)* putative null mutant, in which no staining was observed. White arrows indicate the ventral nerve cord and nerve ring. Grey arrowhead indicates the vulval opening. a, b, Ventral nerve cord staining in the vulval region. Dorsal nerve cord staining is less intense and not visible in this focal plane. c, d, Nerve ring staining.

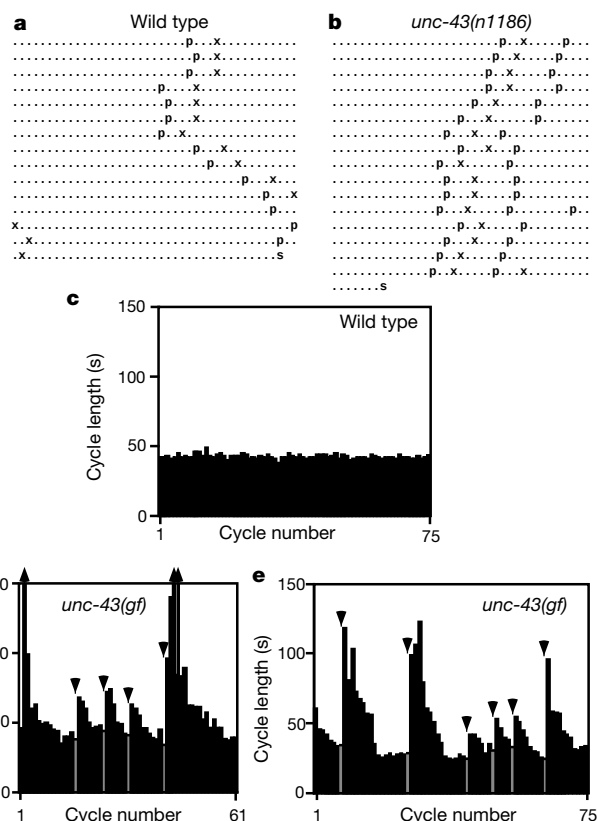


Figure 3 *unc-43* regulates a behavioural clock. Ethograms¹⁶ of wild-type (a) and *unc-43(n1186)* (b) defecation motor programs. Each character represents 1 s; p, posterior body contraction; x, expulsion (that is, EMC); s, record end. c, A record of sequential defecation cycle periods in wild type. y-axis shows time, each bar represents one cycle. d, e, Records of sequential defecation-cycle periods of two different *unc-43(n498)* animals. Up arrows indicate cycles longer than 150 s. Down arrows and grey bars indicate cycles ending with an EMC.

enteric muscles or their motor neurons⁹. Conversely, *unc-43(lf)* mutations mildly and nonspecifically suppressed EMC defects caused by mutations in other genes (Fig. 4b), indicating that *unc-43(lf)* may cause hyperexcited enteric muscle. EMC defects are readily measured and were used to investigate the effects of varying gene dosage of wild-type and *unc-43(gf)* subunits (Fig. 4a). The percentage of defecation cycles that included an EMC (%EMC) was very low in *n498gf* homozygotes. The %EMC was similar when the dose of *n498* was reduced (*n498gf/null*) but was less severe when *n498* was co-expressed with a wild-type allele (*n498gf/+*). These observations are consistent with the multimeric structure of CaMKII and the reported phosphorylation of Ca²⁺-activated subunits by adjacent Ca²⁺-activated subunits⁶. As with enteric muscles, egg-laying muscle excitation was reduced by *unc-43(gf)* and enhanced by *unc-43(lf)*, as assayed by egg-laying rate and the stage of eggs laid (Fig. 4c, d). Pharmacological experiments have indicated that *unc-43(gf)* may function postsynaptically in the egg-laying muscles⁹, although additional presynaptic functions are possible.

unc-43 mutations affect locomotion in a manner suggesting an amalgam of neuronal and muscular defects. The body posture of undisturbed *unc-43(lf)* animals tends to be slightly flaccid and has a flattened wave form, suggesting reduced muscle excitation⁹. Touch-stimulated *unc-43(lf)* animals have a weak shrinker phenotype, which results from all four body-wall muscle quadrants contracting simultaneously and is associated with defects in the DD and VD inhibitory motor neurons^{18,19}. Also, *unc-43(lf)* animals often kink their bodies when backing, a phenotype associated with excitatory motor neuron defects²¹. Finally, *unc-43(lf)* animals are variably

Table 1 *unc-43(gf)* effects on body muscle and locomotion

Genotype	Relative body length (%)	Radial movement (mm min ⁻¹)	
		Stimulated	Unstimulated
Wild type	100 ± 0.6	9.0 ± 0.9	1.56 ± 0.15
<i>unc-43(gf)</i>	84.7 ± 0.7	2.5 ± 0.2	0.25 ± 0.05
<i>tax-4(ks11); unc-43(gf)</i>	97.6 ± 0.8	5.4 ± 0.5	0.12 ± 0.01
<i>tax-4(ks11)</i>	98.8 ± 0.8	6.8 ± 0.6	1.09 ± 0.12

The short body length of *unc-43(gf)* ($P < 0.0001$) was suppressed by *tax-4(ks11)* to a wild-type level ($P = 0.26$). Relative body length was calculated as a percentage of the mean length of wild type. For touch-stimulated radial movement, *tax-4(ks11)* also significantly suppressed the *unc-43(gf)* locomotion phenotype ($P < 0.0001$), and *tax-4(ks28)* and *tax-4(p678)* showed similar suppression (not shown). *unc-43(n498)* and *tax-4(ks11)-unc-43(n498)* behaved similarly in assays of spontaneous movement (unstimulated radial movement). The standard error of the mean is given for each experiment.

convulsive, often spontaneously contracting and relaxing their body-wall muscle in brief repeating bursts reminiscent of seizures. In contrast, *unc-43(gf)* (Table 1) causes tonically contracted body-wall muscle, resulting in a shorter body length and lethargic locomotion, a phenotype that may be reciprocal to the convulsive activity caused by *unc-43(lf)*. The lethargy and body-wall muscle phenotypes of *unc-43(gf)* are distinguished by two criteria. First, some mutants with greater body-wall muscle hypercontraction move faster than *unc-43(gf)*⁹. Second, *tax-4* mutations²² suppress the body length defect caused by *unc-43(gf)*, but *tax-4; unc-43(gf)* double mutant animals are still severely lethargic (Table 1). Specifically, *tax-4* mutations suppressed *unc-43(gf)* movement in response to prodding, a measure of coordinated locomotion, but did not suppress the *unc-43(gf)* movement defect when animals were not prodded. Based on these results, we propose that *unc-43* separately regulates body-wall muscle excitation, spontaneous activity levels and perhaps other aspects of locomotion. The complexity of *unc-43* locomotion phenotypes is consistent with the broad expression of CaMKII and its potentially diverse functions in neurons and muscles.

CaMKII presumably exerts its physiological influence by phosphorylating effector proteins. Some of these effectors are probably ion channels, although little direct *in vivo* evidence supports this assertion. To test this possibility, we constructed double mutants between *unc-43(gf)* and loss-of-function mutations in known ion-channel genes, reasoning that loss of an UNC-43 effector in a tissue might abolish the *unc-43(gf)* phenotype in that tissue. We found that *unc-103(lf)* mutations strongly suppressed the enteric muscle defect but not other defects caused by *unc-43(gf)* (Fig. 5a). This

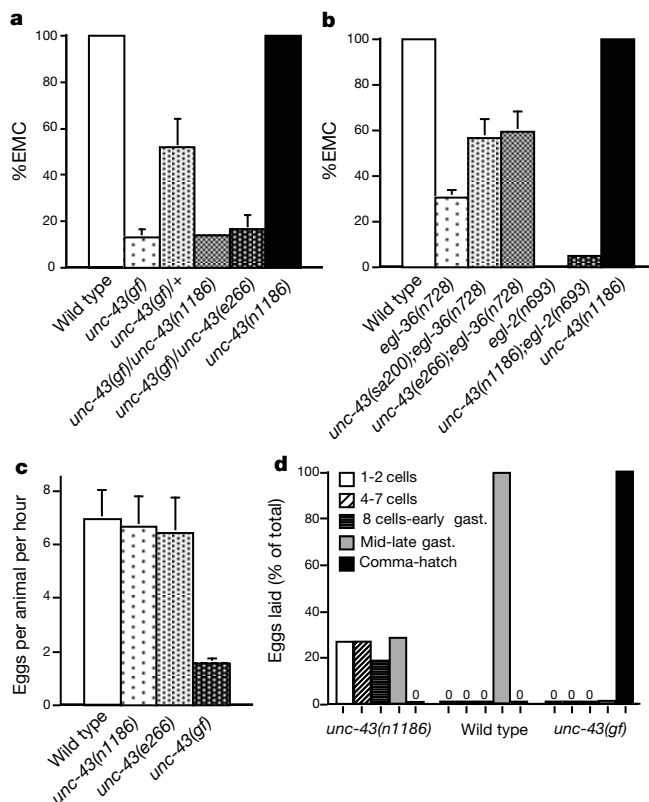


Figure 4 Egg-laying and enteric muscle defects in *unc-43* mutants. **a**, *unc-43(gf)* dosage analysis EMC. **b**, *unc-43(lf)* effects on enteric muscles. *unc-43(lf)* mildly suppresses EMC defects caused by mutations in *egl-36* ($P = 0.005$) and *egl-2* ($P < 0.0001$). *e266* and *sa200* alone have 100% EMC (not shown). **c**, Egg-laying rate. Because eggs were retained in the uterus, *unc-43(gf)* laid fewer eggs per hour ($P < 0.0001$) than wild type or *unc-43(n1186)*. **d**, Stage of embryos. Compared with wild type, *unc-43(n498)* laid fewer advanced-stage embryos ($P < 0.0001$), and *unc-43(lf)* laid earlier stage embryos ($P < 0.0001$). Gastr, gastrulation.

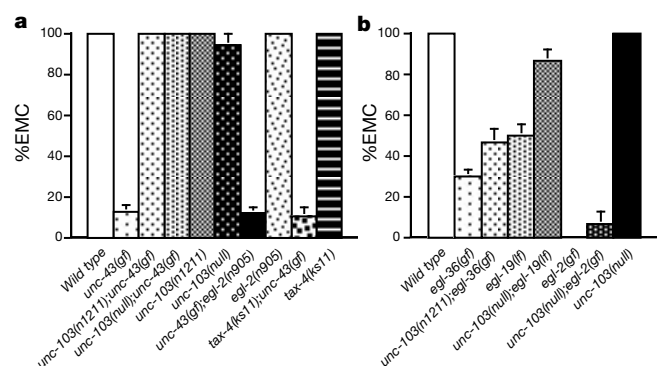


Figure 5 *unc-103* interacts genetically with *unc-43*. **a**, EMC defect suppression. *unc-103(n1213)* is a null mutation (D.J.R. and J.H.T., unpublished data) and does not suppress other characterized phenotypes of *unc-43(gf)*. Mutations in several other muscle excitation genes did not significantly suppress the *unc-43(gf)* EMC defect (see Methods). **b**, Suppression is specific. Mutations in other genes causing a range of severity of EMC phenotypes are shown singly and in double mutant combinations with *unc-103(lf)*.

suppression was highly specific: *unc-103(lf)* did not strongly suppress enteric muscle defects caused by mutations in other genes (Fig. 5b); and *lf* mutations in other genes controlling enteric muscle excitation did not suppress *unc-43(gf)* (see Methods). The strength and specificity of this interaction suggests that UNC-43 regulates UNC-103 directly, rather than functioning in parallel. *unc-103* encodes the nematode homologue of the human ERG voltage-gated K⁺ channel (D.J.R. *et al.*, unpublished data), a member of the *eag* family that regulates cardiac muscle excitability^{23–25}. A simple interpretation of these results is that the UNC-103 K⁺ channel is activated by CaMKII phosphorylation, and that the resulting potentiation of K⁺ efflux accounts for the *unc-43(gf)* EMC defect. Loss of *unc-103* function alone causes relatively subtle phenotypes (Fig. 5; and data not shown), but *unc-43(gf)* causes marked *unc-103*-dependent enteric muscle excitation defects. This indicates that potentiation of the channel by *unc-43(gf)* may have effects out of proportion to the wild-type contribution of *unc-103* to cell excitability.

Its interaction with UNC-103 implicates *eag* family K⁺ channels as *unc-43* targets *in vivo*. In *Drosophila*, loss of *eag* function and expression of a peptide antagonist of CaMKII produced similar learning deficits, and CaMKII can phosphorylate the *eag* protein *in vitro*²⁶. The interaction between *unc-103* and *unc-43* provides much stronger *in vivo* evidence for such an interaction. It also suggests that CaMKII can exert a major regulatory effect through a single target, rather than acting on a complex set of gene products.

The scope of CaMKII functions *in vivo* was previously unclear. Deleting or activating the mouse CaMKII α causes very mild phenotypes, restricted largely to deficits in long-term potentiation and learning^{2,3,27,28}. In contrast, *unc-43(lf)* and *gf* mutations cause a variety of obvious defects in neuronal and muscular function. This difference might be due to the presence of four CaMKII genes in mice, whereas *C. elegans* has only one. The four mouse genes might be partially redundant, or each gene might mediate specific processes, all of which are mediated by *unc-43* in *C. elegans*. In addition to the variety of behavioural defects described here, *unc-43* mutations also perturb synaptic localization of a GFP-tagged GLR-1 glutamate receptor²⁹. Thus, the nematode CaMKII probably mediates cell excitability through both modulation of channel activity and modelling of the synapse. □

Methods

Molecular analysis

unc-43 three-factor map data: *daf-14* (4/35) *unc-43* (4/35) *mec-3* (27/35) *dpy-20*. Cosmid sequence was provided by the nematode genome sequencing consortium¹⁰. Further sequencing was performed using the ABI dye-terminator system. Sequence was analysed using DNA Strider version 1.2, the NCBI BLAST service and ClustalW version 1.4. cDNAs were isolated using a probe from the second *unc-43* exon. The size of intron 1 was approximated using PCR amplification of genomic DNA with primers designed from the cDNA sequence of exons 1 and 2.

Antibody staining

Animals were fixed and permeabilized as described³⁰. Fixed animals were incubated with 4A11 antibody diluted 1:25 in AbA (1X PBS, 1% BSA, 0.5% Triton X-100, 0.05% Na Azide and 1 mM EDTA) at 4 °C with rocking for 16 h, then washed for several hours at 4 °C with several changes of AbB (AbA with 0.1% BSA). Secondary antibody was anti-mouse-FITC (Boehringer Mannheim) diluted 1:200 in AbA, incubated and washed as above. DAPI was added during slide mounting. *lin-15(n765)* multivulva animals were included as internal controls for staining of *n1186* and *e755* mutants. Photographs were taken at 1000 $\times$ magnification using ASA400 Ektachrome film.

Behavioural assays

EMCs were recorded as described⁹. Egg-laying rate and stage were assayed as described³. For spontaneous locomotion assays, animals were placed in the centre of a 10-cm plate with a confluent lawn of *Escherichia coli* OP50 at least three days old, and the origin was marked. The plates were incubated for 15 min at 23 °C, then chilled briefly to stop movement. The radial distance that each animal moved from the origin was measured to the nearest 0.5 mm. Touch-stimulated assays were similar except that each animal was tapped on the tail 2–5 times with a platinum pick, and the assay was for 1 min. Statistical analyses used the Mann–Whitney U-test, except for egg-laying rate and stage, which were

analysed using the χ^2 test. Error bars represent standard error of the mean, except for egg-laying rate, for which they represent standard deviation. *lf* mutations in the muscle excitation genes *unc-93*, *egl-2*, *egl-23*, *tax-2*, *tax-4*, *unc-58* and *unc-105* did not suppress the *unc-43(gf)* EMC and other defects.

Measurement of animal length

Animals were staged at late L4 and photographs were taken at 50 $\times$ magnification. Images were digitally scanned and printed at approximately 200 $\times$ magnification. Animals were measured from the tip of their nose to the posterior end of the intestine using a SCALEX plan wheel. Fifteen animals were measured for each genotype except *ks11* (ten animals).

Received 16 July; accepted 9 September 1999.

- Schulman, H. & Hanson, P. I. Multifunctional Ca²⁺/calmodulin-dependent protein kinase. *Neurochem. Res.* **18**, 65–77.
- Silva, A. J., Stevens, C. F., Tonegawa, S. & Wang, Y. Deficient hippocampal long-term potentiation in alpha-calcium-calmodulin kinase II mutant mice. *Science* **257**, 201–206 (1992).
- Silva, A. J., Paylor, R., Wehner, J. M. & Tonegawa, S. Impaired spatial learning in alpha-calcium-calmodulin kinase II mutant mice. *Science* **257**, 206–211 (1992).
- Kanaseki, T., Ikeuchi, Y., Sugiura, H. & Yamauchi, T. Structural features of Ca²⁺/calmodulin-dependent protein kinase II revealed by electron microscopy. *J. Cell Biol.* **115**, 1049–1060 (1991).
- De Koninck, P. & Schulman, H. Sensitivity of CaM kinase II to the frequency of Ca²⁺ oscillations. *Science* **279**, 227–230 (1998).
- Hanson, P. I., Meyer, T., Stryer, L. & Schulman, H. Dual role of calmodulin in autophosphorylation of multifunctional CaM kinase may underlie decoding of calcium signals. *Neuron* **12**, 943–956 (1994).
- Schulman, H. The multifunctional Ca²⁺/calmodulin-dependent protein kinases. *Curr. Opin. Cell Biol.* **5**, 247–253 (1993).
- Park, E. C. & Horvitz, H. R. Mutations with dominant effects on the behavior and morphology of the nematode *Caenorhabditis elegans*. *Genetics* **113**, 821–852 (1986).
- Reiner, D. J., Weinshenker, D. & Thomas, J. H. Analysis of dominant mutations affecting muscle excitation in *Caenorhabditis elegans*. *Genetics* **113**, 821–852 (1986).
- The *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012–2108 (1998).
- Brocke, L., Srinivasan, M. & Schulman, H. Developmental and regional expression of multifunctional Ca²⁺/calmodulin-dependent protein kinase isoforms in rat brain. *J. Neurosci.* **15**, 6797–6808 (1995).
- Griffith, L. C. & Greenspan, R. J. The diversity of calcium/calmodulin-dependent protein kinase II isoforms in *Drosophila* is generated by alternative splicing of a single gene. *J. Neurochem.* **61**, 1534–1537 (1993).
- Gutay, B. & Griffith, L. C. Functional heterogeneity of alternatively spliced isoforms of *Drosophila* Ca²⁺/calmodulin-dependent protein kinase II. *J. Neurochem.* **66**, 1282–1288 (1996).
- Cho, K. O. *et al.* The alpha subunit of type II Ca²⁺/calmodulin-dependent protein kinase is highly conserved in *Drosophila*. *Neuron* **7**, 439–450 (1991).
- White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Phil. Trans. R. Soc. Lond. B* **314**, 1–340 (1986).
- Liu, D. W. & Thomas, J. H. Regulation of a periodic motor program in *C. elegans*. *J. Neurosci.* **14**, 1953–1962 (1994).
- Thomas, J. H. Genetic analysis of defecation in *Caenorhabditis elegans*. *Genetics* **124**, 855–872 (1990).
- McIntire, S. L., Jorgensen, E., Kaplan, J. & Horvitz, H. R. The GABAergic nervous system of *Caenorhabditis elegans*. *Nature* **364**, 337–337 (1993).
- McIntire, S. L., Jorgensen, E. & Horvitz, H. R. Genes required for GABA function in *Caenorhabditis elegans*. *Nature* **364**, 334–337 (1993).
- Dal Santo, P., Logan, M. A., Chisholm, A. D. & Jorgensen, E. The inositol trisphosphate receptor regulates a 50 second behavioral rhythm in *C. elegans*. *Cell* **98**, 757–767 (1999).
- Hedgecock, E. M., Culotti, J. G. & Hall, D. H. The *unc-5*, *unc-6*, and *unc-40* genes guide circumferential migrations of pioneer axons and mesodermal cells on the epidermis in *C. elegans*. *Neuron* **4**, 61–84 (1990).
- Komatsu, H., Mori, I., Rhee, J. S., Akaike, N. & Ohshima, Y. Mutations in a cyclic nucleotide-gated channel lead to abnormal thermosensation and chemosensation in *C. elegans*. *Neuron* **17**, 707–718 (1996).
- Warmke, J. W. & Ganetzky, B. A family of potassium channel genes related to *eag* in *Drosophila* and mammals. *Proc. Natl Acad. Sci. USA* **91**, 3438–3442 (1994).
- Curran, M. E. *et al.* A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. *Cell* **80**, 795–803 (1995).
- Sanguinetti, M. C., Jiang, C., Curran, M. E. & Keating, M. A mechanistic link between an inherited and an acquired cardiac arrhythmia: HERG encodes the IK, potassium channel. *Cell* **81**, 299–307 (1995).
- Griffith, L. C., Wang, J., Zhong, Y., Wu, C. F. & Greenspan, R. J. Calcium/calmodulin-dependent protein kinase II and potassium channel subunit *eag* similarly affect plasticity in *Drosophila*. *Proc. Natl Acad. Sci. USA* **91**, 10044–10048 (1994).
- Mayford, M., Wang, J., Kandel, E. R. & O'Dell, T. J. CaMKII regulates the frequency-response function of hippocampal synapses for the production of both LTD and LTP. *Cell* **81**, 891–904 (1995).
- Bach, M. E., Hawkins, R. D., Osman, M., Kandel, E. R. & Mayford, M. Impairment of spatial but not contextual memory in CaMKII mutant mice with a selective loss of hippocampal LTP in the range of the theta frequency. *Cell* **81**, 905–915 (1995).
- Rongo, C. & Kaplan, J. M. CaMKII regulates the density of central glutamatergic synapses *in vivo*. *Nature* **402**, 195–199 (1999).
- Miller, D. M. & Shakes, D. C. Immunofluorescence microscopy. *Methods Cell Biol.* **48**, 365–394 (1995).
- Johnstone, D. B., Wei, A., Butler, A., Salkoff, L. & Thomas, J. H. Behavioral defects in *C. elegans egl-36* mutants result from potassium channels shifted in voltage-dependence of activation. *Neuron* **19**, 151–164 (1997).

Acknowledgements

We thank I. Mori for strains and discussion of *tax-4*; T. Inoue for *daf-14* cloning; J. Knapp for isolating *sa402*; K. Iwasaki, D. Johnstone, D. Liu, E. Malone, D. Weinshenker and the

rest of our lab for discussions; G. Garriga and D. Johnstone for reading the manuscript; L. Griffith for 4A11 antibody; Y. Kohara for cDNA clones; A. Coulson for cosmid clones; H. R. Horvitz for *unc-43* mutants; and B. Meyer for hospitality during the writing of this manuscript. Some strains used in this work were provided by the *Caenorhabditis* Genetics Center, which is funded by the NIH Center for Research Resources (NCRR). The work was funded by an NIH grant to J.H.T., and E.M.N. was supported by the Cancer Research Fund of the Damon Runyon-Walter Winchell Foundation.

Correspondence and requests for materials should be addressed to J.H.T. (e-mail: jht@genetics.washington.edu). Sequence is deposited as Genbank AF180735.

APC^{Cdc20} promotes exit from mitosis by destroying the anaphase inhibitor Pds1 and cyclin Clb5

Masaki Shirayama, Attila Tóth, Marta Gálová & Kim Nasmyth

Research Institute of Molecular Pathology (IMP), Dr Bohr-Gasse 7, A-1030 Vienna, Austria

Ubiquitin-mediated proteolysis due to the anaphase-promoting complex/cyclosome (APC/C) is essential for separation of sister chromatids, requiring degradation of the anaphase inhibitor Pds1, and for exit from mitosis, requiring inactivation of cyclin B Cdk1 kinases¹. Exit from mitosis in yeast involves accumulation of the cyclin kinase inhibitor Sic1 as well as cyclin proteolysis mediated by APC/C bound by the activating subunit Cdh1/Hct1 (APC^{Cdh1})^{2,3}. Both processes require the Cdc14 phosphatase, whose release from the nucleolus during anaphase causes dephosphorylation and thereby activation of Cdh1 and accumulation of another protein, Sic1 (refs 4–7). We do not know what determines the release of Cdc14 and enables it to promote Cdk1 inactivation, but it is known to be dependent on APC/C bound by Cdc20 (APC^{Cdc20}) (ref. 4). Here we show that APC^{Cdc20} allows activation of Cdc14 and promotes exit from mitosis by mediating proteolysis of Pds1 and the S phase cyclin Clb5 in the yeast *Saccharomyces cerevisiae*. Degradation of Pds1 is necessary for release of Cdc14 from the nucleolus, whereas degradation of Clb5 is crucial if Cdc14 is to overwhelm Cdk1 and activate its foes (Cdh1 and Sic1). Remarkably, cells lacking both Pds1 and Clb5 can proliferate in the complete absence of Cdc20.

Cells lacking Cdc20 arrest in metaphase with unseparated sister chromatids and high levels of cyclin B/Cdk1 kinases. Deletion of the *PDS1* gene permits separation of sister chromatids, but *cdc20 pds1* double-mutant cells arrest in late anaphase with high levels of Clb2/Cdk1 (refs 8–10). The implication is that APC^{Cdc20} is needed to degrade proteins that inhibit inactivation of Cdk1 (ref. 11). Loss of one such protein might allow mitotic exit and thereby permit proliferation of *cdc20 pds1* double mutants. We therefore set out to identify mutations that permit *cdc20* deletion mutants to proliferate if and only if they lack *PDS1*.

We created a yeast strain whose *CDC20* and *PDS1* genes had been deleted and whose proliferation depended on expression of *CDC20* from a weakened galactose-inducible *GAL1-10* promoter (*GALL*)¹² on a centromeric (*CEN*) plasmid carrying the *URA3* gene. Our parental strain arrests in late anaphase when shifted to glucose medium, with high levels of the mitotic cyclins Clb2 and Clb3 and low levels of Sic1 (Fig. 1a). To avoid the need to clone by complementation, we transformed this strain with a library of yeast DNA mutagenized at random with a bacterial transposon¹³ and identified nine recessive mutations that permitted growth in the presence of glucose, all of which fell into a single complementation group. Remarkably, all nine mutants could grow in the presence of 5-fluoroorotic acid; that is, when the *CEN* plasmid containing

GALL-CDC20 and *URA3* was lost. None of the mutants could proliferate in the presence of glucose upon introduction of a wild-type *PDS1* gene. We prepared genomic DNA from three of the suppressor mutations, amplified the mutated genes using the 'vectorette polymerase chain reaction (PCR)' technique¹⁴ and found, unexpectedly, that all three transposons had been inserted in the S phase cyclin gene *CLB5*.

Clb5 is one of six B-type cyclins in *S. cerevisiae*. It accumulates in late G1 along with the related cyclin Clb6. Activation of Clb5/Cdk1 kinase due to degradation of Sic1 triggers the onset of S phase¹⁵. Deletion of *CLB5* by targeted recombination also enabled proliferation of *cdc20 pds1* double mutants and permitted both proteolysis of Clb2 and Clb3 and accumulation of Sic1 (Fig. 1b). However, deleting *CLB1*, *CLB3*, *CLB4* or *CLB6* had no such effect (Fig. 1c). The effect of deleting *CLB2* could not be tested because *pds1 clb2* double mutants are inviable. A failure to degrade Clb5 appears therefore to be the sole reason why *cdc20 pds1* double mutants cannot degrade Clb2 and Clb3, accumulate Sic1 or exit from mitosis. Deletion of *CLB5* suppressed neither the lethality of *cdc20 pds1 cdh1* triple mutants nor that of *apc2 pds1* double-deletion mutants. The activity of APC^{Cdh1} is therefore crucial for the proliferation of *cdc20 pds1 clb5* triple-mutant cells.

To investigate more precisely when Clb5 is degraded during the cell cycle and whether this depends on APC^{Cdc20}, we analysed strains whose Clb5 was tagged with multiple myc epitopes. In wild-type cells, Clb5 persists within nuclei until shortly before the metaphase-to-anaphase transition, whereupon it suddenly disappears (Fig. 1d, wild type). Degradation of Clb5 takes place earlier than that of Clb2, which only disappears during anaphase⁹. In cells lacking both *CDC20* and *PDS1*, Clb5 and associated Cdk1 kinase persisted, along with Clb2 and Clb3, well after separation of sister chromatids (Fig. 1a). These cells arrested in late anaphase with high levels of Clb5–myc9 within their nuclei (Fig. 1d). Clb5 proteolysis during mitosis therefore resembles that of Pds1 in both its timing and dependence on APC^{Cdc20} (refs 2, 16). The degradation of Clb5 does not depend on Cdh1 at the metaphase-to-anaphase transition (Fig. 1e), consistent with the observation that Clb5 levels are not affected by overexpression of *CDH1* (ref. 3).

Inactivation of Clb2/Cdk1 kinases at the end of mitosis depends on dephosphorylation of Cdh1, Sic1 and Swi5, which is mediated by a highly conserved phosphatase called Cdc14 (refs 4, 5). For much of the cell cycle, Cdc14 is sequestered in the nucleolus owing to its association with the regulatory subunit Cfi1/Net1 (refs 6, 7). The release of Cdc14 from the nucleolus and its accumulation throughout the nucleus (and partially in the cytoplasm) during anaphase is thought to be important for activating APC^{Cdh1} and promoting accumulation of Sic1. To investigate whether APC^{Cdc20} is required for release of Cdc14 from the nucleolus, we used centrifugal elutriation to isolate G1 cells from a *GALL-CDC20* strain expressing a functional myc epitope-tagged Cdc14 protein and incubated them in fresh medium containing glucose. After replicating their DNA and forming buds, most cells arrested with high levels of Clb2, Clb3 and Clb5 and low levels of Sic1 (Clb5/Cdk1 kinase was therefore fully active; Fig. 2a). Indirect immunofluorescence showed that Cdc14–myc18 colocalized with the nucleolar protein Nop2 (Fig. 2a)¹⁷. APC^{Cdc20} is therefore essential for release of Cdc14 from the nucleolus.

To test whether degradation of Pds1 or Clb5 has a role in release of Cdc14 (sister separation *per se* is not required⁷), we repeated this experiment with cells carrying deletions of either *CLB5* or *PDS1*. *cdc20 clb5* mutants failed both to separate sister chromatids and to degrade the mitotic cyclins Clb2 and Clb3. They also failed to accumulate Sic1 and therefore arrested with high levels of Clb2/Cdk1 kinase (Fig. 2b). These results are consistent with reports that Clb2 is stable in cells that overproduce non-degradable Pds1 (refs 18, 19). Furthermore, Cdc14 remained associated with the nucleolus (Fig. 2b). Thus, Pds1 blocks release of Cdc14 from the nucleolus.

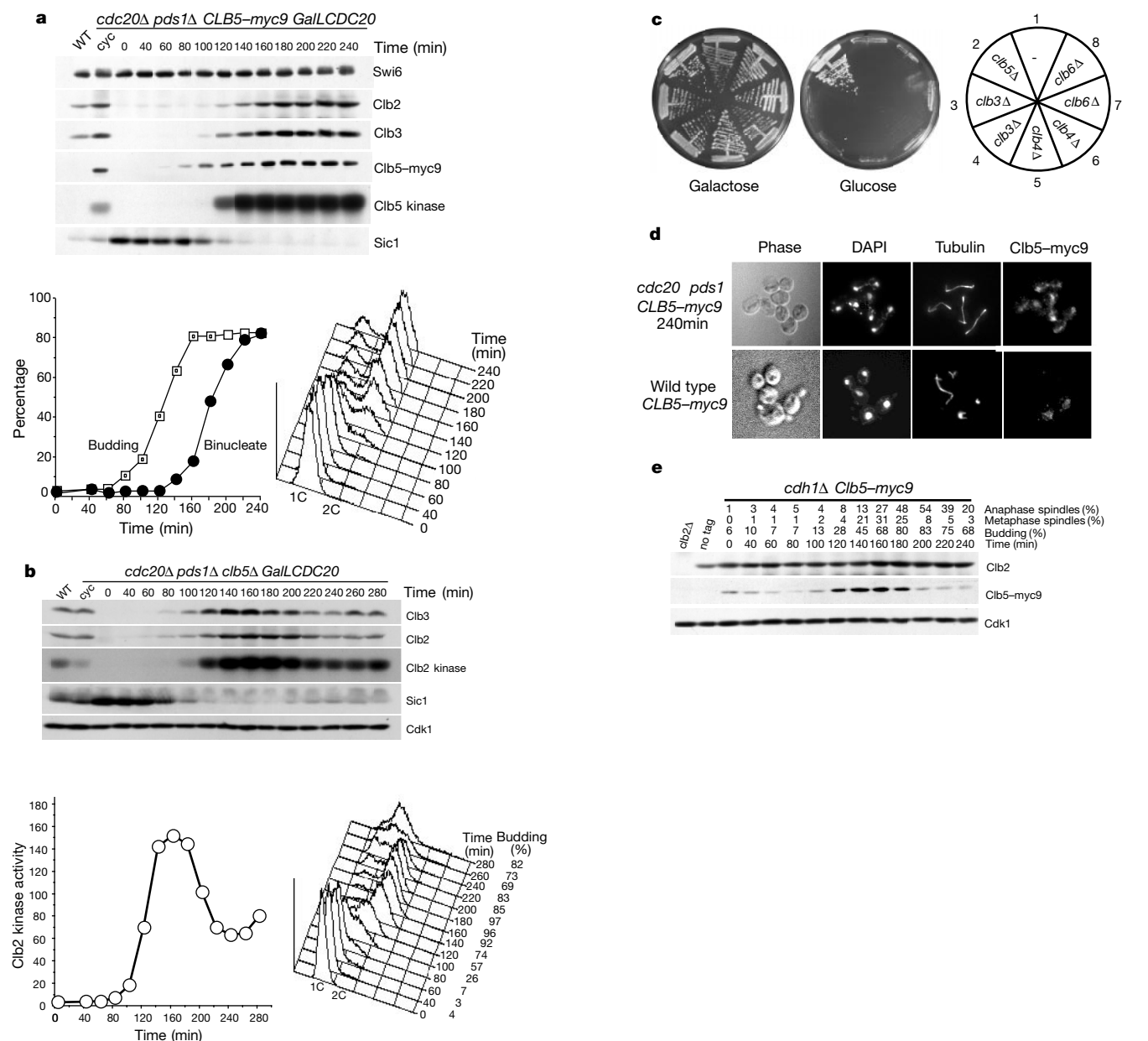


Figure 1 Deletion of *CLB5* bypasses lethality of *cdc20 pds1* null mutants. **a**, Characterization of the mitotic arrest of *cdc20 pds1* mutants. Top, Swi6, Clb2, Clb3, Clb5-myc9 and Sic1 proteins and Clb5-myc9-associated histone H1 kinase activity; bottom right, DNA content and, bottom left, percentage of budding and binucleates of strain K8351 (*cdc20Δ pds1Δ CLB5-myc9*YpGALLCDC20) in synchronous culture after release from elutriation into glucose medium at 25 °C. *CDC20* expression was repressed by incubating cells in raffinose medium for 45 min before elutriation. WT, wild type; cyc, cycling culture of strain K8351. **b**, Deletion of *PDS1* and *CLB5* causes *Cdc20*-independent Clb2/Cdk1 inactivation. Top, Clb3, Clb2, Sic1 and Cdk1 proteins and Clb2-associated histone H1 kinase activity; bottom left, quantification of Clb2/Cdk1 activity; and, bottom right, DNA content of synchronous culture of strain K8352 (*cdc20Δ pds1Δ clb5Δ* YpGALLCDC20) in glucose medium from elutriation. **c**, Only deletion of *CLB5* suppresses lethality of *cdc20 pds1* mutants among B-type cyclin mutations. 1, strain K8356 (*cdc20Δ pds1Δ* YpGALLCDC20); 2, K8357 (*cdc20Δ pds1Δ clb5Δ*); 3 and 4, K8358 and K8359

It is unclear whether Pds1 blocks Clb2's degradation by inhibiting its partner Esp1 (refs 18, 19), because *esp1-1* mutants can, with at most a modest delay, release Cdc14 from the nucleolus⁷ and inactivate Clb2/Cdk1 (ref. 20) despite having failed to separate sister chromatids²¹. Furthermore, *esp1-1* mutants have little or no defect in cyclin proteolysis in the absence of the spindle checkpoint

(*cdc20Δ pds1Δ clb3Δ*YpGALLCDC20); 5 and 6, K8360 and K8361 (*cdc20Δ pds1Δ clb4Δ*YpGALLCDC20); 7 and 8, K8362 and K8363 (*cdc20Δ pds1Δ clb6Δ*YpGALLCDC20). All strains were incubated at 25 °C for 3 days. *cdc20Δ pds1Δ clb1Δ* cells containing YpGALLCDC20 (K8364) are not viable on glucose plate (data not shown). We could not test whether *clb2Δ* can suppress lethality of *cdc20Δ pds1Δ* mutants as *pds1Δ clb2Δ* is synthetic lethal at 25 °C. *cdc20Δ pds1Δ clb5Δ* triple-deletion mutants (K8357) do not contain YpGALLCDC20 plasmid. **d**, Detection of Clb5-myc9 by immunofluorescence in *cdc20 pds1* mutant cells in the arrest at 240 min after release into glucose medium from elutriation or wild-type cells in cycling. Clb5-myc9, staining using mouse anti-myc antibodies; tubulin, staining using rat anti-tubulin antibodies; DAPI (4'-6-diamidino-2-phenylindole) used to stain DNA. **e**, Degradation of Clb5 does not depend on Cdh1. Clb2, Clb5-myc9 and Cdk1 proteins; percentages of budding, metaphase spindles and anaphase spindles of synchronous culture of *cdh1* mutant cells. Small unbudded G1 cells of strain K8418 (*cdh1Δ CLB5-myc9*) were obtained by elutriation and released into glucose medium at 25 °C.

genes *mad2* and *bub2* (Fig. 2c). This indicates that Pds1 might block release of Cdc14 by acting on a target other than Esp1. However, the phenotype of null alleles of *ESP1* will need to be established to settle this issue. In *cdc20 pds1* cells, Cdc14 colocalized with Nop1 until metaphase but spread to the rest of the nucleus (and partly to the cytoplasm) at around the time of nuclear division (Fig. 3), as found

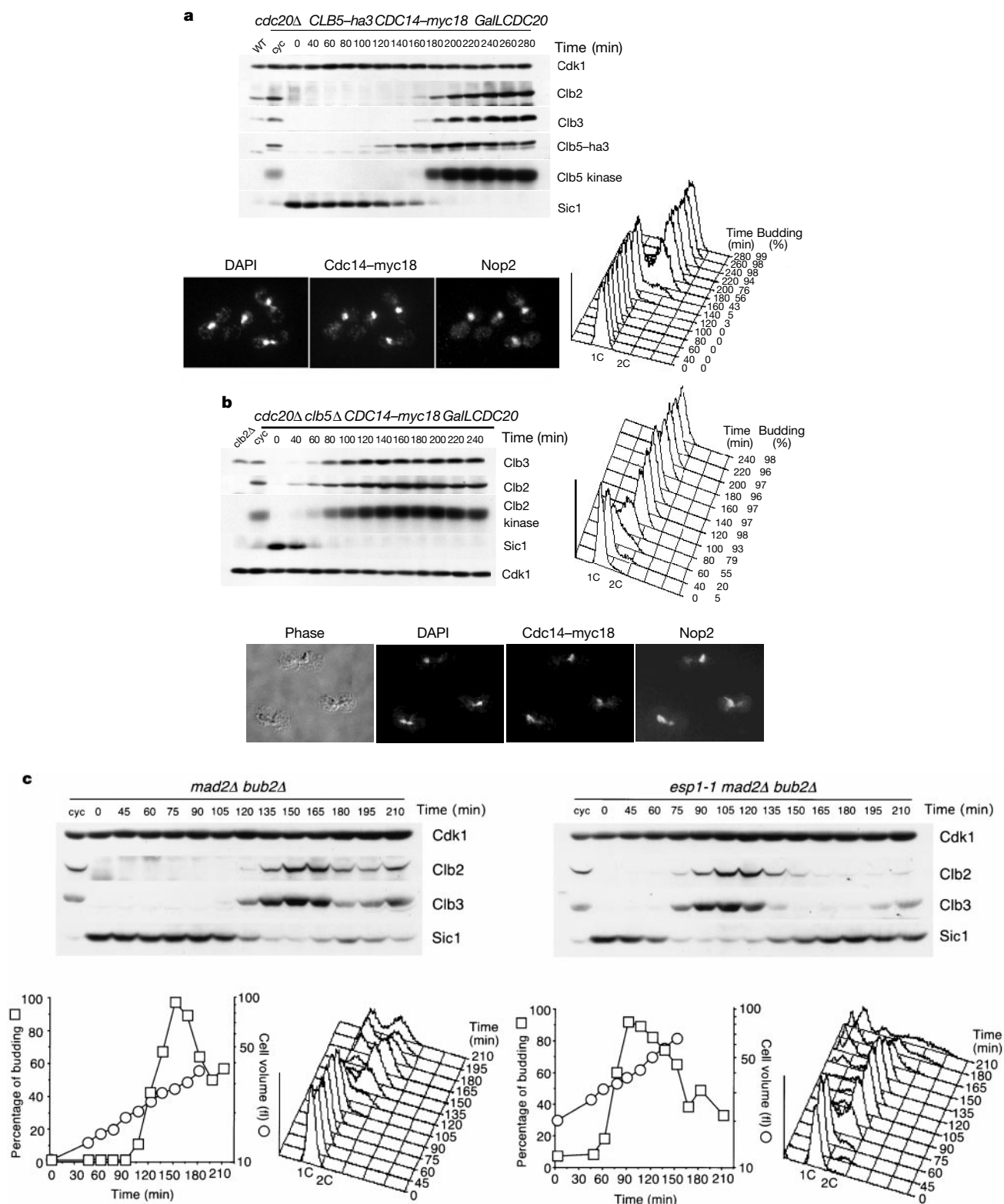


Figure 2 Characterization of *cdc20* and *cdc20 clb5* mutants. **a**, Cdc20 is required for delocalization of Cdc14. Top, Cdk1, Clb2, Clb3, Clb5-ha3 and Sic1 proteins and Clb5-ha3-associated histone H1 kinase activity; bottom right, flow-cytometric DNA analysis of synchronous culture of *cdc20* mutant cells. Small unbudded G1 cells of strain K8353 (*cdc20Δ CLB5-HA3 CDC14-myc18 YCpGALLCDC20*) were obtained by elutriation and released into fresh glucose medium at 25 °C as in Fig. 1. Bottom left, detection of Cdc14-myc18 and Nop2 proteins by immunofluorescence of *cdc20* mutant cells in the arrest at 280 min after release. Cdc14-myc18 staining using mouse anti-myc antibodies; Nop2 staining using rabbit anti-Nop2 antibodies¹⁷. **b**, Pds1 inhibits not only sister separation but also delocalization of Cdc14. Top left, Clb3, Clb2, Sic1 and Cdk1 proteins and Clb2-

associated histone H1 kinase activity; top right, DNA content of strain K8354 (*cdc20Δ clb5Δ CDC14-myc18 YCpGALLCDC20*) in synchronous culture, obtained and released as in Fig. 1. Bottom, detection of Cdc14-myc18 and Nop2 by immunofluorescence in *cdc20 clb5* cells in the arrest at 240 min after release. **c**, *esp1-1 mad2Δ bub2Δ* mutants have little defect in cyclin degradation. Cdk1, Clb2, Clb3 and Sic1 proteins; percentages of budding and cell volume of synchronous cultures of *mad2Δ bub2Δ* mutant (left) or *esp1-1 mad2Δ bub2Δ* mutant (right) cells. Small unbudded G1 cells of strain K7764 (*mad2Δ bub2Δ*) or K8469 (*esp1-1 mad2Δ bub2Δ*) were obtained by elutriation and released into glucose medium at 37 °C.

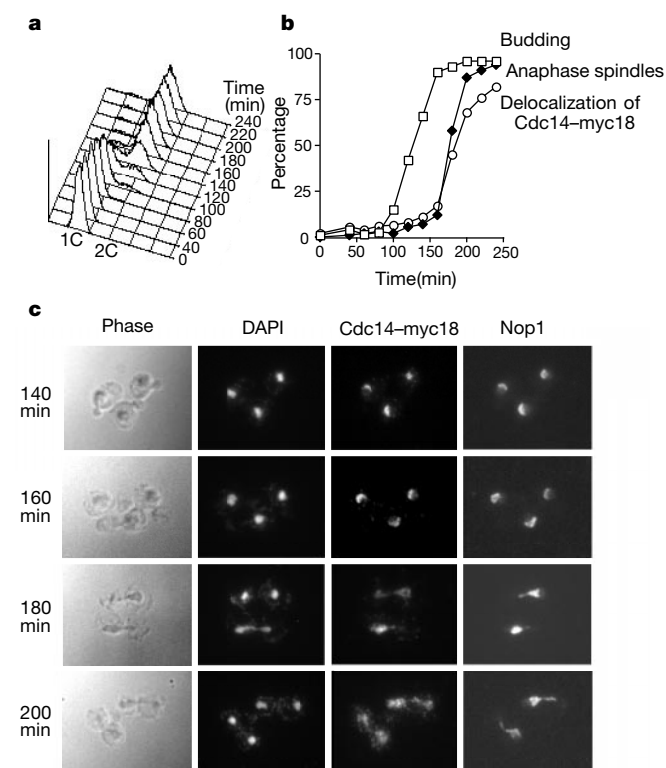


Figure 3 Cdc14 delocalizes into nucleus in *cdc20 pds1* mutants. **a**, DNA content; **b**, percentage of budding (squares), anaphase spindles (diamonds) and delocalization of Cdc14-myc18 (circles) of synchronous culture of strain K8355 (*cdc20Δpds1ΔCDC14-myc18* YCpGALLCDC20) from elutriation. **c**, Subcellular localization of Cdc14-myc18 during the cell cycle. Samples were taken and fixed at the indicated time points from synchronous culture. Cdc14-myc18, staining using rabbit anti-myc antibodies; Nop1, staining using mouse anti-Nop1 antibodies.

in wild-type cells^{6,7}. We conclude that degradation of Pds1 but not of Clb5 is required for release of Cdc14 from the nucleolus. Our data indicate that *cdc20 pds1* double mutants fail to activate Cdh1 and accumulate Sic1 despite having released Cdc14 from the nucleolus.

Our data imply that proteolysis of Clb5 and Pds1 by APC^{Cdc20} at the metaphase-to-anaphase transition mediates Cdk1 inactivation by two mechanisms (Fig. 4). Proteolysis of Pds1 permits release of Cdc14 from the nucleolus, whereas proteolysis of Clb5 renders this phosphatase capable of activating Sic1 and Cdh1. Clb5/Cdk1 readily phosphorylates both Sic1 and Cdh1 *in vitro* and causes Sic1 to be unstable and Cdh1 to be inactive^{22–25}. We suggest that it performs this task *in vivo* at a rate that exceeds the dephosphorylation of Sic1 and Cdh1 by Cdc14 and that this is why Clb5/Cdk1 must be inactivated by a mechanism that does not require either Sic1 or Cdh1. Nevertheless, we cannot exclude the possibility that Clb5/Cdk1 also directly inhibits Cdc14 activity in a Cfi1/Net1-independent way.

To ensure that DNA replication origins fire only once during the cell cycle, the eukaryotic cell cycle consists of two main phases: a low-Cdk1 state permissive to pre-replication complex assembly, and a high-Cdk1 state that triggers replication from these complexes and inhibits their *de novo* assembly¹¹. Irreversible transitions between these two self-maintaining steady states are defined by Clb5/Cdk1 activity: activation of Clb5/Cdk1 marks the onset of the high-Cdk1 state, whereas its proteolysis by APC^{Cdc20} is essential for activating the Cdk1 antagonists (Cdh1 and Sic1) needed to generate a stable low-Cdk1 state. Destruction of an S phase cyclin by APC^{Cdc20} might also be important in Cdk1 inactivation in animal cells. Clb5's behaviour resembles that of cyclin A, which is involved in DNA replication and whose destruction seems to be essential for mitotic exit^{26–29}.

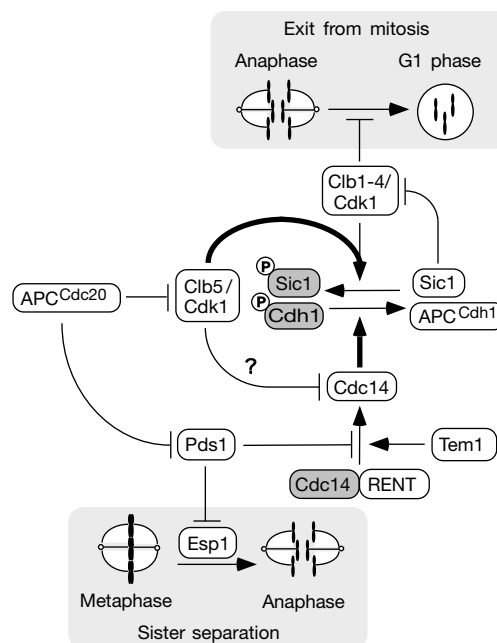


Figure 4 APC^{Cdc20} coordinates 'sister separation' and 'exit from mitosis' during the cell cycle. APC^{Cdc20} targets Pds1 for destruction to promote sister separation and emancipation of Cdc14 from nucleolar RENT complex. Cdc14 can dephosphorylate and activate Cdh1 and Sic1 for exit from mitosis upon Clb5 degradation by APC^{Cdc20}. Cdc14 dissociation from RENT complex might be separately controlled by a signal transduction pathway containing Tem1 GTPase and Cdc15 protein kinase^{6,7}.

Coordinating proteolysis of B-type cyclins with that of anaphase inhibitors like Pds1 might help to ensure that Cdk1 is only inactivated when cells segregate their chromosomes. Pds1 blocks not only sister separation but also release of Cdc14 from the nucleolus. Thus Pds1's destruction alone contributes to the coordination of Cdk1 inactivation with sister separation. However, cells completely lacking the APC^{Cdc20} regulatory system can still coordinate chromosome segregation and cytokinesis in the absence of Clb5 and Pds1 (Fig. 1b). Cdc14's release from the nucleolus remains regulated, perhaps by the Tem1 GTPase pathway^{6,7}, and this presumably delays Cdk1 inactivation until after sister chromatid separation is triggered on time by a still-unknown Pds1-independent mechanism. □

Methods

We carried out genetic screening using a *cdc20Δpds1ΔGALLCDC20* strain (K8356), in the presence of glucose, at 23 °C to avoid complications with the temperature-sensitive phenotype of a *pds1* deletion. Cell volume was measured as described¹⁵. The *CDC14* gene was tagged at its carboxy terminus with 18 copies of a myc epitope by one-step PCR methods³⁰. All other methods are as described⁹.

Received 7 July; accepted 6 September 1999.

- Morgan, D. O. Regulation of the APC and the exit from mitosis. *Nature Cell Biol.* **1**, E47–E53 (1999).
- Visintin, R., Prinz, S. & Amon, A. *CDC20* and *CDH1*: a family of substrate-specific activators of APC-dependent proteolysis. *Science* **278**, 460–463 (1997).
- Schwab, M., Lutum, A. S. & Seufert, W. Yeast Hct1 is a regulator of Clb2 cyclin proteolysis. *Cell* **90**, 683–693 (1997).
- Visintin, R. *et al.* The phosphatase Cdc14 triggers mitotic exit by reversal of Cdk-dependent phosphorylation. *Mol. Cell* **2**, 709–718 (1998).
- Jaspersen, S. L., Charles, J. F. & Morgan, D. O. Inhibitory phosphorylation of the APC regulator Hct1 is controlled by the kinase Cdc28 and the phosphatase Cdc14. *Curr. Biol.* **9**, 227–236 (1999).
- Shou, W. *et al.* Exit from mitosis is triggered by Tem1-dependent release of the protein phosphatase Cdc14 from nucleolar RENT complex. *Cell* **97**, 233–244 (1999).
- Visintin, R., Hwang, E. S. & Amon, A. Cfi1 prevents premature exit from mitosis by anchoring Cdc14 phosphatase in the nucleolus. *Nature* **398**, 818–823 (1999).
- Yamamoto, A., Guacci, V. & Koshland, D. Pds1p, an inhibitor of anaphase in budding yeast, plays a critical role in the APC and checkpoint pathway(s). *J. Cell Biol.* **133**, 99–110 (1996).
- Shirayama, M., Zachariae, W., Ciosk, R. & Nasmyth, K. The Polo-like kinase Cdc5p and the WD-repeat protein Cdc20p/Fizzy are regulators and substrates of the anaphase promoting complex in *Saccharomyces cerevisiae*. *EMBO J.* **17**, 1336–1349 (1998).
- Lim, H. H., Goh, P. Y. & Surana, U. Cdc20 is essential for the cyclosome-mediated proteolysis of both Pds1 and Clb2 during M phase in budding yeast. *Curr. Biol.* **8**, 231–234 (1998).

11. Novák, B. *et al.* Finishing the cell cycle. *J. Theor. Biol.* **199**, 223–233 (1999).
12. Mumberg, D., Mueller, R. & Funk, M. Regulatable promoters of *Saccharomyces cerevisiae*: comparison of transcriptional activity and their use for heterologous expression. *Nucleic Acids Res.* **22**, 5767–5768 (1994).
13. Burns, N. *et al.* Large-scale analysis of gene expression, protein localization, and gene disruption in *Saccharomyces cerevisiae*. *Genes Dev.* **8**, 1087–1105 (1994).
14. Riley, J. *et al.* A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res.* **18**, 2887–2890 (1990).
15. Schwob, E., Böhm, T., Mendenhall, M. D. & Nasmyth, K. The B-type cyclin kinase inhibitor p40^{SIC1} controls the G1 to S transition in *S. cerevisiae*. *Cell* **79**, 233–244 (1994).
16. Cohen-Fix, O., Peters, J.-M., Kirschner, M. W. & Koshland, D. Anaphase initiation in *Saccharomyces cerevisiae* is controlled by the APC-dependent degradation of the anaphase inhibitor Pds1p. *Genes Dev.* **10**, 3081–3093 (1996).
17. de Beus, E., Brockenbrough, J. S., Hong, B. & Aris, J. P. Yeast *NOP2* encodes an essential nucleolar protein with homology to a human proliferation marker. *J. Cell Biol.* **127**, 1799–1813 (1994).
18. Tinker-Kulberg, R. L. & Morgan, D. O. Pds1 and Esp1 control both anaphase and mitotic exit in normal cells and after DNA damage. *Genes Dev.* **13**, 1936–1949 (1999).
19. Cohen-Fix, O. & Koshland, D. Pds1p of budding yeast has dual roles: inhibition of anaphase initiation and regulation of mitotic exit. *Genes Dev.* **13**, 1950–1959 (1999).
20. Surana, U. *et al.* Destruction of the CDC28/CLB mitotic kinase is not required for the metaphase to anaphase transition in budding yeast. *EMBO J.* **12**, 1969–1978 (1993).
21. Ciosk, R. *et al.* An Esp1/Pds1 complex regulates loss of sister chromatid cohesion at the metaphase to anaphase transition in yeast. *Cell* **93**, 1067–1076 (1998).
22. Skowyra, D. *et al.* F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. *Cell* **91**, 209–219 (1997).
23. Feldman, R. M., Correll, C. C., Kaplan, K. B. & Deshaies, R. J. A complex of Cdc4p, Skp1p, and Cdc53p/Cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. *Cell* **91**, 221–230 (1997).
24. Verma, R. *et al.* Phosphorylation of Sic1p by G1 CDK required for its degradation and entry into S phase. *Science* **278**, 455–460 (1997).
25. Zachariae, W., Schwab, M., Nasmyth, K. & Seufert, W. Control of cyclin ubiquitination by CDK-regulated binding of Hct1 to the anaphase promoting complex. *Science* **282**, 1721–1724 (1998).
26. Epstein, C. B. & Cross, F. R. *CLB5*: a novel B cyclin from budding yeast with a role in S phase. *Genes Dev.* **6**, 1695–1706 (1992).
27. Schwob, E. & Nasmyth, K. *CLB5* and *CLB6*, a new pair of B cyclins involved in DNA replication in *Saccharomyces cerevisiae*. *Genes Dev.* **7**, 1160–1175 (1993).
28. Sherr, C. J. Mammalian G1 cyclins. *Cell* **73**, 1059–1065 (1993).
29. Sigrist, S., Jacobs, H., Stratmann, R. & Lehner, C. F. Exit from mitosis is regulated by *Drosophila* fizzy and the sequential destruction of cyclins A, B and B3. *EMBO J.* **14**, 4827–4838 (1995).
30. Wach, A. *et al.* Heterologous *HIS3* marker and GFP reporter modules for PCR-targeting in *Saccharomyces cerevisiae*. *Yeast* **13**, 1065–1075 (1997).

Acknowledgements

We thank J. Charles, D. Morgan, A. Amon and E. Schwob for sharing unpublished results; E. Kramer and J.-M. Peters for technical help; A. Amon, J. Aris, L. Huber, L. Guarente, D. Kellogg, and J. Kilmartin for gifts of antibodies; G. Ammerer for the transposon library; and F. Cross, Y. Kikuchi and W. Zachariae for comments on this manuscript.

Correspondence and requests for materials should be addressed to K.N.
(e-mail: nasmyth@nt.imp.univie.ac.at).

The RNA splicing factor hSlu7 is required for correct 3' splice-site choice

Katrin Chua & Robin Reed

Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA

The production of correctly spliced messenger RNA requires two catalytic splicing steps^{1–3}. During step II, exon 1 attacks an adenine-guanine (AG) dinucleotide at the 3' splice site. This AG is usually located between 18 and 40 nucleotides downstream from the branch site, and closer AGs are skipped in favour of AGs located more optimally downstream^{4–6}. At present, little is understood about how the correct AG is distinguished from other AGs. Here we describe a metazoan splicing factor (hSlu7) that is required for selection of the correct AG. In the absence of hSlu7, use of the correct AG is suppressed and incorrect AGs are activated. We investigated this loss of fidelity by analysing spliceosomes assembled in the absence of hSlu7. These studies reveal that

exon 1 is loosely associated with these spliceosomes. Thus, the improperly held exon cannot access the correct AG, but can attack other AGs indiscriminately. We conclude that hSlu7 is required to hold exon 1 tightly within the spliceosome for attack on a prespecified AG.

We have shown that the human homologue of the yeast step II splicing factor Slu7 (refs 7, 8) functions in step II in humans⁹. To determine whether hSlu7 is involved in AG selection, we have now examined 3' splice-site choice in HeLa cell nuclear extracts immunodepleted of hSlu7. Consistent with previous work⁹, a pre-mRNA containing a single AG located 23 nucleotides (nt) from the branch site underwent step I, but not step II, in hSlu7-depleted (Δ hSlu7) extracts (Fig. 1a). As expected^{4–6}, when a duplicate AG was inserted 11 nt from the branch site, this AG was skipped, and the downstream AG was selected in mock-depleted and normal extracts (Fig. 1b, mock; and data not shown). The upstream AG, however, was specifically activated and the normally used (downstream) AG was suppressed in Δ hSlu7 extracts (Fig. 1b, Δ hSlu7). This activation of the upstream AG and suppression of the normal AG is general, as it also occurred in Δ hSlu7 extracts when the two AGs were located at 11 and 18 nt or 9 and 23 nt from the branch site (Fig. 1c; and data not shown). We conclude that upstream AGs that are normally silent are specifically activated in the absence of hSlu7, whereas the normally used AG is specifically suppressed.

To further investigate the basis for the Δ hSlu7-specific activation of normally silent AGs, we examined pre-mRNAs in which the upstream AG is selected in normal splicing extracts. In the first pre-mRNA tested, the two AGs were located 15 and 22 nt from the branch site. In mock extracts, the AG at 15 nt was used efficiently (Fig. 2a, mock). In Δ hSlu7 extracts, step II was inefficient as indicated by the accumulation of exon 1 and the lariat-exon. Significantly, however, the low level of step II products that was detected shows that the normal (upstream) AG is suppressed whereas the downstream AG is specifically activated (Fig. 2a, Δ hSlu7). The same results were observed with pre-mRNAs containing AGs located 17 and 23 or 18 and 24 nt downstream from the branch site (Fig. 2b, c). In all cases, the first AG downstream from the branch site is selected in normal extracts, whereas this AG is specifically suppressed and the normally silent AG is activated in the Δ hSlu7 extracts. This aberrant AG selection is specific to Δ hSlu7 extracts, as it did not occur in extracts immunodepleted of the step II factors hPrp16 (Z. Zhou and R.R., unpublished data) and hPrp18 (ref. 10). As expected, addition of recombinant hSlu7, but not hPrp16, to the hSlu7-depleted extract is sufficient to restore correct AG selection (Fig. 2d).

Together, our data show that aberrant AG selection occurs in Δ hSlu7 extracts whether the duplicate AG is located upstream (Fig. 1) or downstream (Fig. 2) of the normal AG. Activation of incorrect AGs only occurs when AGs are located within a maximal distance of ~30 nt from the branch site (data not shown). Significantly, aberrant AG selection is not random, as the AG that is used in normal splicing extracts is specifically suppressed, and the normally silent AG is activated. Thus, in the absence of hSlu7, the normal AG is recognized in some manner, but cannot be used.

To gain insight into why correct 3' splice sites are suppressed and incorrect AGs are activated, we characterized the spliceosome that assembles in extracts depleted of hSlu7. The early spliceosomal complexes A and B assemble normally in Δ hSlu7 extracts, but spliceosomal complex C, which contains the products of catalytic step I, fails to assemble; instead, a novel complex, the Δ hSlu7 spliceosome, accumulates⁹. To isolate this complex, we used a new method in which spliceosomal complexes are resolved on native gels composed of low-melting-point agarose¹¹. As expected, the C complex that formed in mock and normal extracts contained both splicing intermediates (the lariat-exon 2 and exon 1) (Fig. 3b, lanes 2 and 14)¹². Strikingly, however, the Δ hSlu7 spliceosome contained the lariat-exon 2, as observed previously⁹, but completely

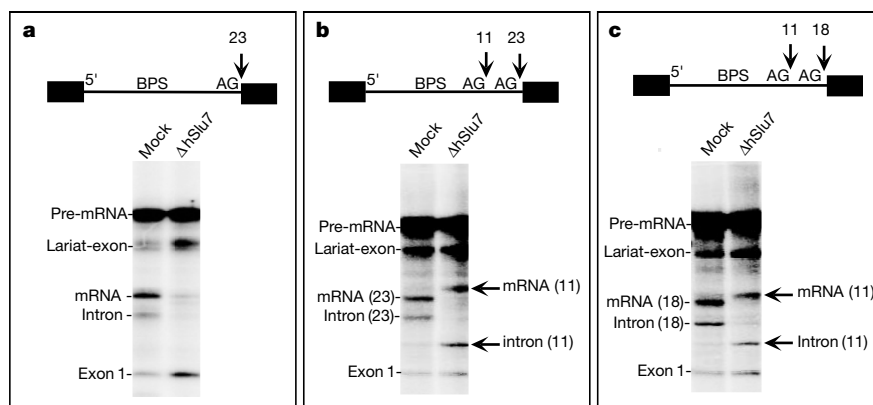


Figure 1 Suppression of the correct AG and activation of aberrant upstream AGs in the absence of hSlu7. Δ hSlu7 and mock-depleted extracts were used to splice AdML (adenovirus major late) pre-mRNA derivatives containing a single AG located 23 nt from the branchpoint sequence (BPS) (a), duplicate AGs located 11 and 23 nt from the BPS (b), or duplicate AGs located 11 and 18 nt from the BPS (c). The same results were obtained

with mock-depleted or normal extracts. Total RNA was fractionated on 8% denaturing polyacrylamide gels. The efficiency of step II is decreased in mock extracts when an AG is inserted upstream of the normal AG, but this decrease was not observed with other pre-mRNAs.

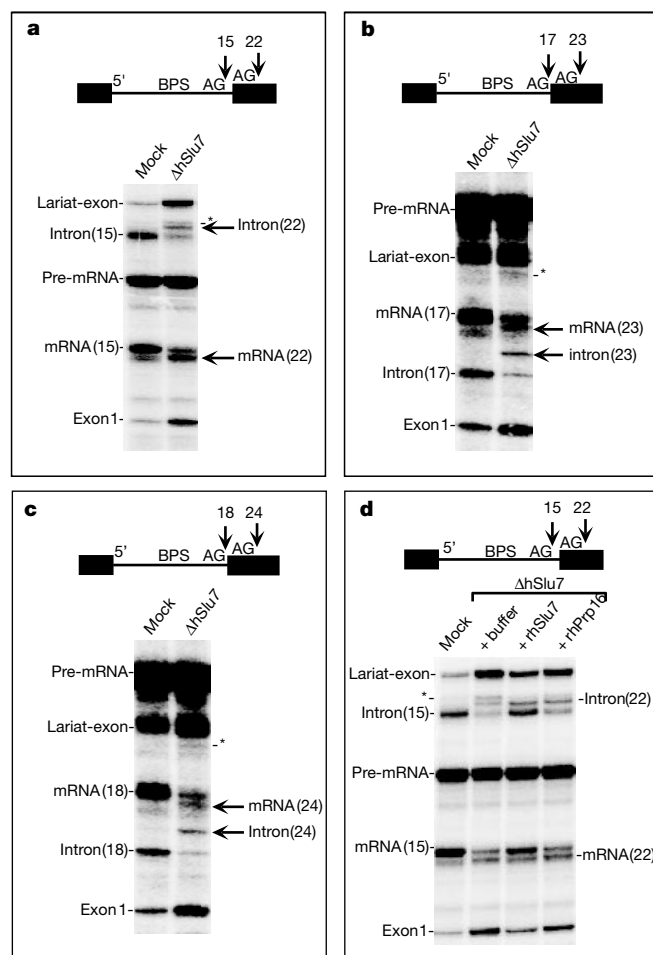


Figure 2 Suppression of the correct AG and activation of aberrant downstream AGs in the absence of hSlu7. Δ hSlu7 and mock-depleted extracts were used to splice pre-mRNAs containing duplicate AGs located 15 and 22 nt (a), 17 and 23 nt (b), or 18 and 24 nt from the BPS (c). d, Reconstitution of proper AG selection to Δ hSlu7 extracts with recombinant hSlu7 (rhSlu7). Asterisk denotes a breakdown product of the lariat-exon-2 intermediate. RNA was fractionated on 13% (a, d) or 8% (b, c) denaturing polyacrylamide gels.

lacked exon 1 (Fig. 3b, lane 10). The absence of exon 1 had not been observed before because of the difficulty of isolating significant quantities of the mutant complexes from polyacrylamide gels⁹.

The polyanion heparin disrupts weak interactions and is usually added to splicing reactions before fractionation on native gels (for example, see Fig. 3a)¹². Exon 1 was not detected in Δ hSlu7 spliceosomes fractionated on native gels in the presence of heparin; however, when the splicing reactions were fractionated on a native agarose gel in the absence of heparin, the Δ hSlu7 spliceosome migrated with the same mobility as the normal C complex (Fig. 3c). Moreover, exon 1 was detected in the Δ hSlu7 spliceosome (Fig. 3d). We conclude that exon 1 is associated with the Δ hSlu7 spliceosome, but is more loosely held than in the normal spliceosome. Moreover, the large difference in mobility of the Δ hSlu7 spliceosome compared with the normal C complex indicates that there may be a major change in conformation correlated with the presence or absence of the exon.

The observation that exon 1 is loosely associated with the Δ hSlu7 spliceosome might explain why step II is blocked in pre-mRNAs containing a single AG. Because step II does occur in pre-mRNAs containing two AGs, it is possible that exon 1 is held more tightly in these Δ hSlu7 spliceosomes. To test this possibility, the pre-mRNA containing AGs at 11 and 23 nt from the branch site was assembled into spliceosomes in normal or Δ hSlu7 extracts. When these spliceosomes were fractionated on native agarose gels in the presence of heparin, both the C complex and the Δ hSlu7 spliceosome had the same migration behaviour as observed on pre-mRNAs containing a single AG (Fig. 3a; and data not shown)⁹. Whereas the normal C complex contained both the lariat-intermediate and exon 1, the Δ hSlu7 spliceosome completely lacked exon 1 (Fig. 4, lanes 1 and 2, and legend). Thus, exon 1 is loosely associated with Δ hSlu7 spliceosomes even when this exon attacks aberrant AGs.

Our data indicate that, in the absence of hSlu7, exon 1 is loosely held within the spliceosome and attack at the correct AG is suppressed. If other AGs are located nearby, these incorrect AGs are attacked by the improperly held exon. The normal AG is specifically suppressed in the absence of hSlu7, which indicates that this AG is both pre-specified and sequestered in some manner. One model for how this occurs is that the active site of the spliceosome is configured around the normal AG before exon 1 is incorporated into this active site. Alternatively, the correct AG may be sequestered by a proofreading factor. In this case, the active site of the spliceosome may be configured around the 3'-OH of exon 1,

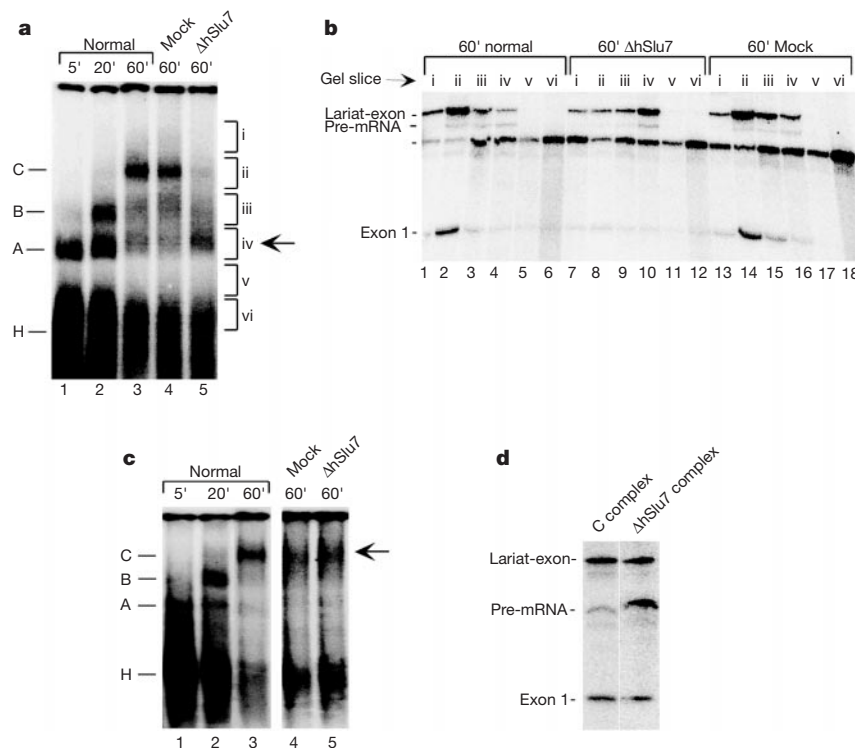


Figure 3 The attacking exon is not properly associated with Δ hSlu7 spliceosomes. **a, b**, Exon 1 is absent from Δ hSlu7 spliceosomes treated with heparin. **a**, A pre-mRNA containing a GG substitution of the AG at the 3' splice site was incubated in normal, mock-depleted or Δ hSlu7 extracts and fractionated on a native agarose gel. The same results were obtained using a pre-mRNA containing an AG at the 3' splice site, but less of the normal C complex accumulates because it is converted to spliced products (data not shown). The arrow indicates the position of the Δ hSlu7 spliceosome. **b**, RNA was isolated

from gel slices shown in **a** and run on a 13% denaturing polyacrylamide gel. i–vi correspond to gel slices from the 60' normal, Δ hSlu7 and mock lanes. Asterisk denotes the breakdown product of the lariat-exon. **c, d**, Exon 1 is associated with Δ hSlu7 spliceosomes isolated under gentle conditions. **c**, Splicing reactions lacking heparin were separated on a native agarose gel. Arrow points to the C complex (lane 4) or Δ hSlu7 complex (lane 5). **d**, RNA was isolated from the C complex and Δ hSlu7 spliceosomes shown in **(c)** and separated on a 13% denaturing gel.

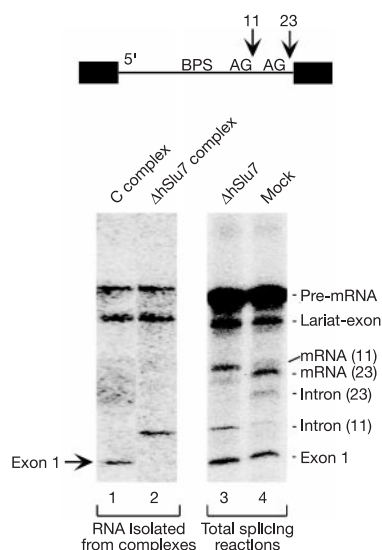


Figure 4 The improperly held exon 1 attacks aberrant AGs. Spliceosomes were assembled on 11AG/23AG pre-mRNA in mock or Δ hSlu7 extracts and fractionated on a native agarose gel. RNA was isolated from the C complex (lane 1) or the Δ hSlu7 spliceosome (lane 2) and fractionated on an 8% denaturing gel. Total RNA isolated from aliquots of the same splicing reactions (lanes 3 and 4) shows that exon 1 is present in the Δ hSlu7 splicing reaction. The lariat-intron was detected in the gel slice containing the Δ hSlu7 spliceosome. This lariat may be inefficiently released from the Δ hSlu7 spliceosome or may be present in an intron complex that comigrates with the Δ hSlu7 spliceosome.

which is thus potentially reactive with any AG, and hSlu7 is required for attack at the correct AG.

Regardless of the details, our data show that normal AG selection occurs in two definable stages. In the first stage, the correct AG is identified but sequestered, and this process does not require hSlu7. In the second stage, the sequestered AG is exposed, and exon 1 is directed to this AG; this stage requires hSlu7. Division of AG selection into at least two steps would enhance the fidelity of step II because the AG is recognized multiple times before it is used^{4,13}.

In yeast, Slu7 is thought to function in recruiting distant AGs^{7,8}, partly on the basis that an upstream AG is chosen over a downstream AG in a Slu7 mutant strain. Whether the correct AG is suppressed and aberrant AGs are activated was not investigated in the yeast studies. Thus, it will be interesting to determine whether yeast Slu7 functions similarly to hSlu7 in AG choice and also in holding exon 1 properly in the spliceosome. Our conclusions about hSlu7 are consistent with Slu7 being synthetically lethal with U5 snRNP in yeast. U5 snRNP is thought to participate in tethering exon 1 to the spliceosome and aligning the two exons for ligation^{16–25}; however, other work indicates that the interactions of U5 snRNP with the exons are unlikely to be sufficient for these processes^{4,21,22,24,26}. We have now identified hSlu7 as a critical factor for holding exon 1 properly within the spliceosome and for directing this exon to the correct AG. □

Methods

Plasmids

Plasmids encoding pre-mRNAs were constructed by inserting oligonucleotides into the *Hind*III and *Acl*I sites of pADML²⁷ to replace the sequences from the branchpoint sequence

to the AG dinucleotide. Beginning with the branch-site adenosine, the sequences in this region are: 11AG/23AG, 5'-aCUUUUUUUCAGUUUUUUUUCAG-3'; 23AG, 5'-aCUUUUUUUUCGGUUUUUUUUCAG-3'; 11AG/18AG⁵, 5'-aCUUUUUUUCCA-GUUUGCAG-3'; 15AG/22AG⁵, 5'-aCUUUUUUUUUUCAGUUUGCAG-3'; 17AG/23AG, 5'-aCUUUUUUUUUUUCAGUUUGCAG-3'; 18AG/24AG, 5'-aCUUUUUUUU-CUUUUUUUCAGUUUCAG-3'.

Splicing reactions and analysis

hSlu7-depleted or mock-depleted nuclear extracts were prepared by rotating with hSlu7 or pre-immune antibodies, immobilized on protein A sepharose⁹. Pre-mRNAs were incubated in these extracts under standard splicing conditions for 1 h. Total RNA was prepared and fractionated on 8% or 13% denaturing polyacrylamide gels. For native gels, spliceosomal complexes were assembled in 25 µl reaction mixtures, and 3 µl of 6.5 mg ml⁻¹ heparin in 50% Ficoll dye or 5 µl of 50% Ficoll dye was added. Aliquots (4 µl) were fractionated at room temperature on horizontal minigels (8 cm × 7 cm × 0.25 cm) composed of 2% low-melting-point agarose in 200 mM Tris, 200 mM glycine running buffer. Gels were run at 60 V for 10 min and then at 100 V for 2 h. RNA was isolated from slices of the gel by melting at 80 °C in 200 mM Tris (pH 8), 25 mM EDTA, 300 mM NaCl, 2% SDS. Samples were treated with proteinase K, extracted with phenol, precipitated, resuspended in loading dye and run on 8% or 13% denaturing polyacrylamide gels.

Received 16 July; accepted 18 August 1999.

- Moore, M. J., Query, C. C. & Sharp, P. A. in *The RNA World* (eds Gesterland, R. F. & Atkins, J. F.) 303–357 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1993).
- Burge, C. B., Tuschl, T. H. & Sharp, P. A. in *The RNA World* 2nd edn (eds Gesteland, R. F., Cech, T. R. & Atkins, J. F.) 525–560 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1999).
- Staley, J. P. & Guthrie, C. Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92**, 315–326 (1998).
- Umen, J. G. & Guthrie, C. The second catalytic step of pre-mRNA splicing. *RNA* **1**, 869–885 (1995).
- Chiara, M. D., Palandjian, L., Feld Kramer, R. & Reed, R. Evidence that U5 snRNP recognizes the 3' splice site for catalytic step II in mammals. *EMBO J.* **16**, 4746–4759 (1997).
- Smith, C. W., Chu, T. T. & Nadal-Ginard, B. Scanning and competition between AGs are involved in 3' splice site selection in mammalian introns. *Mol. Cell. Biol.* **13**, 4949–4952 (1993).
- Frank, D. & Guthrie, C. An essential splicing factor, SLU7, mediates 3' splice site choice in yeast. *Genes Dev.* **6**, 2112–2124 (1992).
- Brys, A. & Schwer, B. Requirement for SLU7 in yeast pre-mRNA splicing is dictated by the distance between the branchpoint and the 3' splice site. *RNA* **2**, 707–717 (1996).
- Chua, K. & Reed, R. Human step II splicing factor hSlu7 functions in restructuring the spliceosome between the catalytic steps of splicing. *Genes Dev.* **13**, 841–850 (1999).
- Horowitz, D. S. & Krainer, A. R. A human protein required for the second step of pre-mRNA splicing is functionally related to a yeast splicing factor. *Genes Dev.* **11**, 139–151 (1997).

- Das, R. & Reed, R. Resolution of the spliceosomal complex E and the ATP-dependent spliceosomal complexes on native agarose gels. *RNA* (in the press).
- Konarska, M. M. & Sharp, P. A. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. *Cell* **46**, 845–855 (1986).
- Reed, R. & Palandjian, L. in *Eukaryotic mRNA Processing* (eds Hames, B. D. & Glover, D. M.) 103–122 (Oxford Univ. Press, Oxford, 1997).
- Frank, D., Patterson, B. & Guthrie, C. Synthetic lethal mutations suggest interactions between U5 small nuclear RNA and four proteins required for the second step of splicing. *Mol. Cell. Biol.* **12**, 5197–5205 (1992).
- Umen, J. G. & Guthrie, C. Mutagenesis of the yeast gene PRP8 reveals domains governing the specificity and fidelity of 3' splice site selection. *Genetics* **143**, 723–739 (1996).
- Newman, A. J. & Norman, C. U5 snRNA interacts with exon sequences at 5' and 3' splice sites. *Cell* **68**, 743–754 (1992).
- Newman, A. J., Teigelkamp, S. & Beggs, J. D. snRNA interactions at 5' and 3' splice sites monitored by photoactivated crosslinking in yeast spliceosomes. *RNA* **1**, 968–980 (1995).
- Wyatt, J. R., Sontheimer, E. J. & Steitz, J. A. Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. *Genes Dev.* **6**, 2542–2553 (1992).
- Cortes, J. J., Sontheimer, E. J., Seiwert, S. D. & Steitz, J. A. Mutations in the conserved loop of human U5 snRNA generate use of novel cryptic 5' splice sites in vivo. *EMBO J.* **12**, 5181–5189 (1993).
- Sontheimer, E. J. & Steitz, J. A. The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science* **262**, 1989–1996 (1993).
- O'Keefe, R. T., Norman, C. & Newman, A. J. The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast. *Cell* **86**, 679–689 (1996).
- Teigelkamp, S., Newman, A. J. & Beggs, J. D. Extensive interactions of PRP8 protein with the 5' and 3' splice sites during splicing suggest a role in stabilization of exon alignment by U5 snRNA. *EMBO J.* **14**, 2602–2612 (1995).
- Umen, J. G. & Guthrie, C. A novel role for a U5 snRNP protein in 3' splice site selection. *Genes Dev.* **9**, 855–868 (1995).
- Newman, A. J. The role of U5 snRNP in pre-mRNA splicing. *EMBO J.* **16**, 5797–5800 (1997).
- O'Keefe, R. T. & Newman, A. J. Functional analysis of the U5 snRNA loop 1 in the second catalytic step of yeast pre-mRNA splicing. *EMBO J.* **17**, 565–574 (1998).
- Ségault, V. *et al.* Conserved loop I of U5 small nuclear RNA is dispensable for both catalytic steps of pre-mRNA splicing in HeLa nuclear extracts. *Mol. Cell. Biol.* **19**, 2782–2790 (1999).
- Gozani, O., Patton, J. G. & Reed, R. A novel set of spliceosome-associated proteins and the essential splicing factor PSF bind stably to pre-mRNA prior to catalytic step II of the splicing reaction. *EMBO J.* **13**, 3356–3367 (1994).

Acknowledgements

We thank K. Clouse, O. Gozani, R. Gaur, B. Graveley, K. Hertel and Z. Zhou for critical comments on the manuscript. This work was supported by an NIH grant to R.R.

Correspondence and requests for materials should be addressed to R.R. (e-mail: reed@hms.harvard.edu).

Looking for a reaction

Monoclonal and polyclonal antibodies, immunoassays, immunoet cetera.

Custom antibody production

Rockland www.rockland.inc.com
Antibodies to your immunogens

Rockland Immunochemicals, based in Gilbertsville, Pennsylvania, produce both polyclonal and monoclonal antibodies to order. Polyclonals can be produced to immunogens provided by the researcher, via immunization with native proteins, GST-fusion proteins, peptides, carbohydrates or any other compound of viral, fungal or microbial origin. Using proprietary immunization methodologies, as little as 50 µg of antigen can be used. Antigens can be supplied to Rockland in soluble form, lyophilized, nitrocellulose membrane or gel slice.

Reader Service No. 100

Antibodies to luciferase

Biogenesis www.biogenesis.co.uk
Antibodies for research and manufacturing

In response to rising demand, Biogenesis, of Poole in Dorset, have raised monoclonal and rabbit polyclonal antibodies to the mono-oxygenase enzyme luciferase. Both antibodies have now been released on general sale and are available in purified Ig format. Activity has been checked against native and recombinant luciferase.

Reader Service No. 101

Smad sampler

Zymed Laboratories www.zymed.com
A pack of antibodies to four Smads

Signalling events leading to transcriptional activation initiated by members of the TGF-β superfamily are known to be mediated by Smad proteins. Biological activities mediated by Smads include cell growth and morphogenesis, development and immune response. Zymed Laboratories is now offering a Smad Antibody Sampler Pack, containing 20-µg aliquots of rabbit polyclonal antibodies to Smad1, Smad2, Smad3 and Smad6. Each antibody is epitope affinity-purified from rabbit antiserum and is suitable for ELISA and western blotting.

Reader Service No. 102



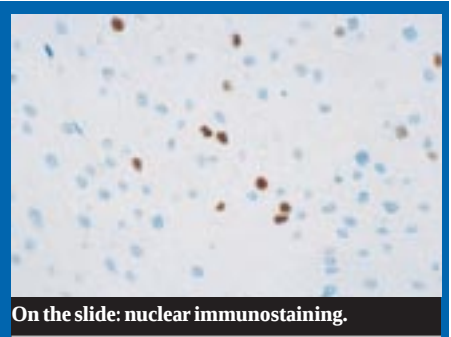
Smad, Smad, Smad world.

Papillomavirus antibody

Dako www.dakousa.com
Monoclonal mouse anti-human papillomavirus (anti-HPV), clone K1 H8

This anti-HPV reacts with a non-conformational internal linear epitope of a major capsid protein of HPV-1, which is expressed among the different HPV subtypes. Anti-HPV is immunoreactive with formalin-fixed, paraffin-embedded human tissues infected with HPV types 6, 11, 16, 18, 31, 33, 42, 51, 52, 56 and 58. Positive immunostaining is confined to the nuclei of infected cells.

Reader Service No. 103



On the slide: nuclear immunostaining.

Immunoassays

EURO/DPC

Assays for the Immulite 2000 analyser

New from EURO/DPC Ltd, three new immunoassays perform quantitative measurement of insulin, albumin and C-peptide, often used as predictors in monitoring diabetic patients. Utilizing monoclonal and polyclonal antibodies, the Immulite Insulin assay can capture and detect insulin from serum or heparinized plasma. Immulite Albumin is designed for the detection and quantification of microalbuminuria at concentrations not detectable by conventional dipstick and acid precipitation tests. And Immulite 2000 C-Peptide monitors C-peptide as an index of insulin secretion.

Reader Service No. 104

Troponin antibodies

Covance Research Products www.babco.com
Antibodies to mammalian cardiac troponin-I

Two new monoclonals from Covance show less than 2% crossreactivity to skeletal muscle troponin, making them suitable as cardiac-specific markers or indicators of myocardial damage via serum ELISA. One antibody detects cardiac troponin-I only when phosphorylated on Ser 24, the other reacts with the C-terminal portion of the molecule and thus detects phosphorylated and unphosphorylated forms of the protein. The antibodies react with human, rat, monkey and baboon cardiac troponin-I.

Reader Service No. 105

Inject mckLH

Pierce Chemical www.piercenet.com
Environment-friendly haemocyanin

Inject Mariculture Keyhole Limpet Haemocyanin (mckLH) is purified from keyhole limpets harvested from the ocean and main-

tained in salt-water tanks. Pierce say that this method is less damaging to marine environments than traditional methods involving repeated harvesting directly from the sea. Also available is Maleimide Activated mckLH, which couples to sulphhydryl-containing peptides to produce a strong immunogenic response for easy antibody production.

Reader Service No. 106

Pig IFN-γ ELISA

Endogen www.endogen.com

An assay sensitive to less than 2 pg per ml to 500 pg per ml of pig interferon-γ

This ELISA is designed to measure IFN-γ in 50-µl samples of serum, plasma or culture supernatant in just three hours. Like Endogen's Pig TNFα ELISA, the Pig IFN-γ kit offers one simple, convenient procedure for all sample types.

Reader Service No. 107

TOPO cloning kits

Invitrogen www.invitrogen.com

Fast cloning and sequencing for PCR products

TOPO Cloning Kits for Sequencing are designed to achieve ligation in 5 minutes, for streamlined sequencing of PCR products. Vectors in the kits include four sequencing primer sites including T3 and T7, which are



TOPO explained.

only 33 base pairs from the PCR product insertion site.

Reader Service No. 108

Akt1 antibodies

Upstate Biotechnology

www.upstatebiotech.com

Antibodies specific to activation states

Akt1 (also known as protein kinase B α and Rac- α) is activated by growth factor stimulation. Factors known to stimulate Akt1 include insulin, platelet-derived growth factor and epidermal growth factor (EGF). For complete activation, Akt1 must be phosphorylated at both Thr 308 and Ser 473 by independent events. Upstate now produce a polyclonal antibody to phospho-Akt1 that can detect Akt1 phosphorylated at position Thr 308, joining the antibody, already on the market, that detects phospho-Akt Ser 473.

Reader Service No. 109

Activated resins

Sterogene Bioseparations www.sterogene.com

Resins for heavy-duty separations

Actigel ALD and Actigel Superflow are stable activated resins with high coupling capacities for proteins, antibodies (up to 50 mg per ml) and peptides, at 90–98% coupling efficiency. Used as an immunoaffinity support, the binding capacity is 2–3 times higher than with 'oriented' (hydrazide) or CNBr-coupling. Actigel resins can be autoclaved or sanitized with NaOH.

Reader Service No. 110

Custom antibodies

BioSource International www.lifescrreen.com

Antibodies to order

BSI provides a complete monoclonal production service, with EIA monitoring at each step, samples sent at regular intervals (serum, subclone supernatant, ascites) for evaluation and optional zonal electrophoretic analysis of the antibody. BioSource have extensive experi-

ence of using various carriers (in addition to KLH and ovalbumin) and immunization protocols designed for molecules with low antigenic potential. Antibodies can be produced in several different species, and other services available include cell line storage.

Reader Service No. 111

J chain monoclonal JC88

BioGenex

www.biogenex.com

Antibody to J chain, a 15-kD acidic protein

B cells secrete J chain at an early stage of differentiation with the expression persisting in those cells destined to produce IgA or IgM. J chain modulates the structure of IgM and influences the IgA dimer assembly. Monoclonal antibody JC88 reacts specifically with J chain in formalin-fixed, paraffin-embedded tissue sections.

Reader Service No. 112

J chain monoclonal 6C2

InnoGenex

www.innogenex.com

Another monoclonal antibody to J chain

Antibody 6C2 (IgG1/ κ) is reactive with human, mouse and chicken J chain. The immunogen used for production of this antibody was human J chain. The monoclonal is appropriate for use in immunohistochemistry, ELISA and western blot techniques. It gives excellent staining on routine formalin-fixed, paraffin-embedded tissue sections without the need for pretreatment. Two formats are available, affinity-purified immunoglobulins and ascites preparations.

Reader Service No. 113

These notes are compiled in the Nature office from information provided by the manufacturers. For more details, fill in the reader service card bound in the journal.

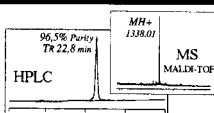
ADVERTISEMENTS

Biomolecular Synthesis oligos, synthetic peptides and polyclonal antibodies

Contact Severn Biotech Ltd
Tel: 01562 825 286
Fax: 01562 825 284
email: info@severnbiotech.com
www.severnbiotech.com

READER ENQUIRY NO. 28

CUSTOM PEPTIDES



- Up to 30 aa
- HPLC & MS data

Purity	>80%	>95%	>99%
Amount	15mg	10mg	

info@genosphere-biotech.com

Tel: +33 1 42 71 70 21 • Fax: +33 1 42 71 70 23

GENOSPHERE BIOTECHNOLOGIES

READER ENQUIRY NO. 24

HIGH QUALITY PEPTIDE SYNTHESIS

- 'Standard' and 'non-standard'
- Fast delivery
- Competitive prices
- Also advisory services

Pepsyn Ltd

Contact: Dr J A Smith, Chairman, Pepsyn Ltd,
School of Biological Sciences, Life Science Building,
The University of Liverpool L69 7ZB, UK
Tel: 0151 794 4329 Fax: 0151 794 4349
e-mail: J.A.Smith@liverpool.ac.uk
Website: www.pepsyn.co.uk

READER ENQUIRY NO. 22

RT activity

the best assays

www.cavidi.com



READER ENQUIRY NO. 29

Diagnostic Reagents

Pefachrome®-Series: chromogenic and fluorogenic substrates

Pefabloc®-Series:
serine proteinase inhibitors

Purified snake venom proteins
(e.g. Protac®, Ecarin, Batroxobin)

PENTAPHARM LTD

CH 4002 Basel Fax: +41 61 319 96 19
<http://www.pentapharm.ch>

READER ENQUIRY NO. 6

NEWCASTLE UNIVERSITY
MOLECULAR BIOLOGY UNIT



PROTEIN SEQUENCING

Membrane-bound, solid or solution samples

£12 per cycle

5 cycles minimum - no set up charge

For more information please contact Dr. Joe Gray:
Phone/Fax: +44 191 222 8612

Email: JOE.GRAY@newcastle.ac.uk

*DNA sequencing (ABI 377-96KL, 877 Catalyst)
also available at competitive prices*

READER ENQUIRY NO. 6

Customised monoclonal & polyclonal antibodies with guarantee!

BioGenes GmbH
Köpenicker Str. 325 • D-12555 Berlin
Tel. +49-30-65 76 23 96 Fax: - 97
E-mail: BioGenes@Berlin.sireco.net
Internet: <http://www.biogenes.de>

READER ENQUIRY NO. 17

FULL DIAGNOSIS

Bedside observation backed up by molecular insights

p214

UNTANGLED WEB

We show you where to look for more career leads online

p214

FUNDING COUNTS

US initiatives aim to entice medical doctors into research work

p215

DOCTOR IN THE LAB

How mentors pointed one medical graduate in the direction of research

p215

From bench to bedside... research makes the translational transition

How is biological research being translated into clinical practice? The challenges of information management need to be applied urgently.

Few would disagree with the view that clinical medical training in the United Kingdom is among the best in the world and that the UK National Health Service delivers a decent standard of medical care to the maximum number of people. However, according to Sir Colin Dollery, a consultant for SmithKline Beecham (SB), the current system fails to produce enough clinically qualified researchers, perhaps resulting in underachievement by the biomedical sector of the UK economy and unrealized dividends from investment in basic biomedical research.

According to Dollery, SB would like to recruit more researchers to handle the discovery of novel compounds in early clinical development. This has been more difficult for the company to do in the United Kingdom than in the United States. "I think if we had five qualified candidates tomorrow, we'd snap them up," says Dollery. He believes that other big pharmaceutical companies are experiencing the same shortage.

At the CRC Laboratories of Imperial College School of Medicine in London, director Charles Coombes has also had difficulty recruiting clinical researchers. "The problem we are having, and I think this applies to the whole of Europe, is that the entire medical education system is oriented to producing functioning doctors and making sure that they are clinically competent," he says.

The percentage of students choosing the path towards clinical research is small, says Coombes, and there is little incentive to follow it for those with original ideas. "By the time they have gone through all this training, they have had all their creativity and drive knocked out of them," says Coombes. Most of those medical students who do get PhDs do so early in their long medical training. For those on a research track, it is usual to have three or four years at the end of their PhD that must be spent doing clinical work, and by the end of this period they are no longer able to be effective as clinical scientists, mainly because they have not been able to keep up with their research interest, he says.

The problem arising from the fact that PhD work is immediately followed by specialized clinical training is compounded, says Dollery, by the fact that clinical training for a specialism is longer in the United Kingdom than in the United States. The typical US medical graduate does three years in a residency programme and three years in a fellowship, followed by board exams. In Britain, there are three years of general medicine followed by four or five years of specialized training, says Dollery, in which students are rotated around various hospitals. By comparison, Dollery cites his US experience at SB, where physician-researchers who trained at Harvard or Washington University often spent the whole of their specialized training at the Brigham and Women's Hospital in Boston or St Louis University Hospital in Missouri. Having a consistent clinical home allows them to keep their interest and expertise in research at a



Dollery: 'qualified staff are in short supply'.

more sustainable level than in the United Kingdom, according to Dollery.

In Britain, individuals can short-circuit training and become accredited to do research in one area. But few people want to do this research-directed specialized training, comments Coombes, because they will be constrained in career and research options. "It's designed to let people out of as much as two years of the standard training, but at the end they don't get the training certificate. They get narrow permission to do research, and they are not regarded by their peers as being clinically competent," says Coombes.

Coombes sees a need for specific, flexible training for clinical scientists, where they are recognized as being of equal calibre to individuals who have gone through the prolonged route and then work as non-research physicians. On funding, he believes "that

After they have gone through all this training, they have had all their creativity knocked out of them.

there have to be grants committees that are specifically for translational projects", stating that only a small percentage of any grants committee is clinically qualified. The scoring systems used by organizations giving grants for health-related research need to have two components: one for the scientific merits of the application, and a second for its applicability to human disease. These should receive equal recognition.

Molecular momentum

A multidisciplinary philosophy of translational research that seems to be gaining speed and mass is being used at Oxford University's Institute of Molecular Medicine (IMM). In its laboratories, such as those occupied by Andrew McMichael's molecular immunology group, basic scientists and clinical scientists work side by side, often on the same problems. "I think there is pressure to do more of that, and I think that the big medical schools want to develop that idea further," says McMichael.

As well as four clinicians, McMichael's group brings together researchers with graduate degrees in biochemistry, immunology, cell biology, chemistry and physics. "We usually have two new MDs come in a year to do full PhD training," he says. Normally, this takes three years and is full-time research from day one. Most join after a medical degree and maybe four years of clinical medicine internships and residency, says McMichael. Often they will do another year of clinical medicine in internships or residency, as well as consultant training lasting four years.

McMichael hopes that the sequencing of

▶ the human genome will provide a big push for translational research. Although the current system does produce some physicians who are fully qualified clinical researchers with a credible laboratory presence, only a few people succeed. "Those who do are really exceptional people," says McMichael. The big problem for the clinician scientist is the demanding clinical training programme, on top of the difficult enough challenge of beginning a research career.

Mechanistic approaches

In the Nuffield department of clinical medicine at Oxford's IMM, John Bell hopes that what has been done in the past to combat infectious diseases can be repeated with chronic diseases, such as diabetes, asthma and cancer. In the same way that an elucidation of the mechanisms of pathogenesis of hepatitis or pleurisy has led to improved understanding of natural history and improved treatment for these infectious diseases, Bell would like to see an analogous mechanistic understanding of chronic multifactorial diseases. Most diseases have been defined historically by their phenotypes but, as the biological basis of disease is unravelled, says Bell, this knowledge will create a taxonomy based on mechanisms.

According to Bell, this approach will have a great impact on the way students learn medicine. "Up to now every textbook on medicine had a chapter on asthma. However, if it turns out that asthma is ten different diseases, with different pathways involved in each, we'll need a very different textbook."

Bell finds some obvious problems with current medical education. It is still oriented towards traditional teaching, high on factual content. Bell would like to see a greater understanding of biology and disease mechanisms, as well as greater emphasis on populations in terms of disease susceptibility and risk.

One striking thing about clinical practice is the heterogeneity of the patient population, says Bell. "Patients are very different even within the same diagnostic criteria," he says. Doctors are taught to 'lump' rather than

'split', says Bell. In the future, however, he believes they will increasingly be taught to split and individualize the way they think about disease. For example, the first question a student will ask when a patient has high blood pressure will be: is that a defect in salt handling by aldosterone, a defect in catecholamine synthesis or a defect in vascular smooth muscle? Bell believes that up to ten per cent of each common disease can be explained mechanistically through highly penetrant, single-gene defects.

Bell wrote in May that "there is little reason to believe that research on patients will remain moribund for long... In the same way that molecular studies were required to breach some of the principal obstacles preventing a mechanistic understanding of disease, so will a resurgence of clinical translational research be necessary to understand how these molecules function in an intact organism" (*Nature Med.* **5**, 477–478; 1999). And, says Bell, "clinical research is emerging as a translational, iterative process involving bedside observation backed up by molecular insights".

Cyber assistance

At the forefront of this coming mechanistic movement is medical informatics, which Bell points out is not taught to students. In the future, clinical researchers will need to become skilled at handling large amounts of data online. "Large-scale databases with this kind of information are going to be the order of the day, and solutions for handling the data will be mostly software solutions."

One reason for the increased reliance on informatics is the issue of splitting diseases, says Bell. Russ Altman, an associate professor of medicine and computer science at Stanford University in California, agrees, saying that medicine has traditionally classified diseases by a constellation of symptoms, including basic lab tests. Today, there are some extremely high-precision lab tests, some of which are still experimental, but are rapidly making their way towards general use, says Altman. "These will allow you to characterize things at a level of detail that allows you to

make many more distinctions. Things that were once one syndrome are being broken up into five, ten, fifteen different sub-syndromes. Bell and Altman agree that analysing this many variations is beyond the processing capacity of the cerebral cortex.

Two branches of informatics are relevant to this revolution: medical informatics — patient records, patient databases, decision support for physicians; and bioinformatics — management of DNA and protein sequence and structure information. In many places, these two have been separate entities. At Stanford they have not been, and Altman says many people are beginning to realize that they should no longer be separate.

From Altman's perspective, this is a natural marriage. "The one area I would point to is pharmacogenetics, where you're trying to understand how variation in human genes among individuals leads to variations in their response to drugs." As a practising physician, he notes that when a patient is given medication there is always a bit of a gamble as to whether it will work. There are probabilities, but these are not really known, he says. Pharmacogenetic studies correlating different genotypes with different responses will allow us to be much more certain of the response when giving medication.

There is also the problem that the techniques for treating molecular entities remain largely unknown, says Altman. "We are going to need a new round of clinical studies where what you randomize on will have to be based on these markers."

In the first part of the Human Genome Project, researchers are taking advantage of the fact that more than 99 per cent of genetic material in humans is identical, Altman explains. The goal is to get that 99 per cent correct and to know where there is variability. Fully characterizing the population's variability is generally thought of as a separate, second-stage project. Such projects include the SNP Consortium, created by the pharmaceutical industry to share information about single-nucleotide polymorphisms, where individuals differ by a single DNA base.

Altman speculates that funding agencies will want to see studies that include a couple of hundred to a couple of thousand patients whose genes are fully sequenced in regions of high interest, with that information correlated to their clinical responses and clinical picture. "In informatics we love this because we're right in the middle of that correlation. We don't collect the data, but once it's collected we put it in the database, we organize it, link it and help analyse it. So it's a very interesting time for informatics."

Altman's triad

"I have this kind of triad in my head of genotypic, clinical phenotypic and more laboratory phenotypic data," says Altman. "If you've observed an allele in a patient,

Information on the web

- American Medical Informatics Association
www.amia.org/
- International Society for Computational Biology
www.iscb.org/
- Institute of Molecular Medicine
www.imm.ox.ac.uk/
- SmithKline Beecham
www.sb.com
- Glaxo Wellcome
www.glaxowellcome.com

- SNP Database
www.ibt.wustl.edu
- NIH's National Center for Research Resources
www.ncrr.nih.gov/
- Howard Hughes Medical Institute
www.hhmi.org/
- Burroughs Wellcome Fund: Clinical Scientists Awards in Translational Research
www.bwfund.org/programsMain.htm

- American Federation for Medical Research
www.afmr.org/
- Association of American Medical Colleges
www.aamc.org/
- Doris Duke Charitable Foundation: Clinical Scientist Awards
web.ortge.ufl.edu/fyi/v25n13/fyi043.html
- FASEB
www.faseb.org/

and you have also observed clinical features, that's nice. But they're usually noisy data, as clinical data tend to be imprecise." But if that allele is put in some experimental system where one can have control over all the parameters, then, says Altman, you will be able to see the physiological response or behaviour for that particular allele.

That triad should give a pretty good picture of what is going on. "However, this is something that physicians are not being trained to think about or to act upon, and this is the challenge for medical education."

According to Altman, there will have to be some 'backward compatibility'. Some retraining will be necessary, but the onus will be on information-systems people to make this easy. No matter what is done with technology, primary-care physicians are still going to have

just 12 minutes per patient, explains Altman. So information systems will have to take less than one minute to help the physician figure out the major decision points, tests that need to be ordered, and the best treatment decisions. Medical informatics has been stressing physician support, which has been regarded as optional until now because good physicians did not need it — they could manage with their brains. "We are looking at a time when even the best physician, unless extremely specialized, is going to need information systems to help manage all the options. Primary-care physicians will need this because their responsibility is simply too broad," concludes Altman.

Brendan Horton

Brendan Horton is a golfer and freelance science writer.

e-mail: b.horton@naturedc.com

Improving the plight of the physician–scientist in the US

For a host of reasons, interest by physicians in careers in clinical research has been waning over the past two decades. Of particular concern is the area of patient-oriented research, which Nobel laureates Michael Brown and Joseph Goldstein of the University of Texas Southwestern Medical School in Dallas define as that which passes the "handshake test" — when the investigator meets the patient during the study.

Although some would argue about the extent of the problem, most would agree that, even though MDs and PhDs have maintained fairly similar success rates when competing for grants from the US National Insti-



Brown: 'funding at the root of the problem'.

tutes of Health (NIH), the number of applications from MDs has fallen far behind those from PhDs, making them an ever-decreasing fraction of those seeking funding. The NIH Director's Panel on Clinical Research reported a 30 per cent drop in first-time grant applications to the NIH from MDs between 1994 and 1996. Applications from PhDs fell just six per cent.

"There have been spotty attempts here and there to patch the problem," says Brown, "but no systematic attempt to fix it. The underlying problem has got much worse, and that is because of the way clinical departments are funded in this country." The growth of managed care, coupled with reduced levels of reimbursement from government health-insurance programmes, has eroded the financial security of academic health centres, placing greater pressures on clinical departments and physicians to boost revenue by providing clinical services to more patients. It is easy to see why academic pursuits often play second fiddle to clinical practice.

In some medical schools, a career in research is not actively promoted or held in such high regard as it once was. This, together with a paucity of clinical training and career-development programmes, a lack of mentors, and a feeling that clinical research grant applications were not getting a fair deal from NIH study sections (although a restructuring of the study sections is under way), have all contributed to physicians' declining interest in clinical research. The prospect of low pay (compared to clinical practice) over the lengthy training period, at a time when most physicians are saddled with debts of \$60,000–80,000, is also a factor.

Why should we care?

This is not to suggest that the biomedical research enterprise is grinding to a halt because physicians have been turning away from clinical research careers. Basic science is flourishing, and many MDs who opt for a career in research choose basic science over clinical studies, in part because the funding situation appears healthier. Increasingly, basic scientists are turning their attention to clinical problems, says Brown. It has become fashionable for basic biologists to affiliate to



After graduating from medical school, **Brenda Nicholson** (above) was intent on a career in clinical practice. But a three-year fellowship in haematology–oncology at Vanderbilt University Medical Center, Nashville, Tennessee, changed all that. **Now an assistant professor of medicine at the university, Nicholson recently received a five-year clinical associate physician, or CAP, award from NIH to further her patient-oriented research in breast cancer.** The CAP award, part of NIH's effort to

Clinical research in practice

attract young physicians into clinical research, provides 75 per cent of an NIH-capped salary. Nicholson, in return, will be expected to devote 75 per cent of her time to the projects outlined in the award.

Nicholson attended medical school at Wright State University in Dayton, Ohio, and followed this with an internship and residency in internal medicine at the Bowman Gray School of Medicine in Winston-Salem, North Carolina. Both colleges run clinically directed programmes with almost no exposure to research, says Nicholson, "and that's why when I came out of training and went into fellowship I had in mind that [ultimately] I was going to go into private practice.

Exposure really was the thing that stimulated me to want to stay [in academic medicine]." During her fellowship at Vanderbilt, Nicholson says she was given considerable responsibilities and opportunities, as well as a lot of positive reinforcement. **She credits her mentor, David H. Johnson,** deputy director of the Vanderbilt–Ingram Cancer Center and director of medical oncology, with helping to launch her career in academic medicine, and the new chairman of medicine **for actively encouraging young physicians to go after NIH funds.**

Vanderbilt is unusual, she says, in being strong both in basic science and clinically. She has managed to develop

collaborative projects with basic scientists, including a study to test the combination of tamoxifen and the monoclonal antibody Herceptin in patients with oestrogen-responsive tumours and metastatic breast cancer, a *BRCA1* gene therapy trial, and a study to test whether taxol can sensitize cancer cells to radiation therapy.

Now with a CAP award under her belt, which affords physicians the kind of mentoring and protected time needed to develop the skills to become independent investigators, Nicholson is very much on an academic medicine career track. She says: **"It's a career that's exciting, and allows you a balance between patient care and the creativity of doing studies."** **D. G.**

companies or to start their own, to try to develop new therapies. "The danger would be if medical schools drift so far away from science... that the students more and more see the research as irrelevant to medicine as they wish to practise it," he says.

Leon Rosenberg of the department of molecular biology at Princeton University, who was a member of the NIH director's panel, says: "We're not raising this issue because the system is falling apart today... we are concerned about the future." When former NIH director James Wyngaarden first drew attention to the problem 20 years ago in his article "The clinical investigator as an endangered species", Rosenberg says the message largely fell on deaf ears.

There are those, adds Rosenberg, who say that, if MDs do not stay in the world of clinical research, their place will be taken by well-trained PhD scientists. "I think that's a myopic view of the role that these different professionals can bring to a research system. MDs are conditioned by their experience with sick people, and the questions they ask are often formed from their experiences."

Initiatives launched

Both the US Institute of Medicine and the NIH director's panel have published reports and made recommendations for tackling the problem. Interestingly, the Federation of American Societies for Experimental Biology, an organization not usually noted as representing the interests of physicians, held a meeting in June on career opportunities for physician-scientists, and is due to present its report to its directors in December.

Both the published reports cited deficiencies in clinical research training, career development and mentoring. Institutions were found to be struggling to engage and launch physicians in careers in clinical research, with the quality of programmes varying widely. It's hard to train people who are good both in the clinic and at the bench, since the knowledge requirement is so enormous, says David Nathan, professor of medicine and faculty dean

for academic programmes at the Dana-Farber Cancer Institute in Boston, who was also a member of the NIH panel.

The Mayo Clinic in Rochester, Minnesota, operates a training programme supporting 10–12 'residents' each year, providing them with two years'

experience in a clinical research lab. "Programmes like ours were born of a frustration at not being able to train and attract these people into clinical investigative careers," says Gregory Poland, professor of medicine

in the Mayo Vaccine Research Group and associate programme director for the Mayo Clinic's General Clinical Research Center. In September, the clinic received one of the new NIH-funded K30 clinical research curriculum awards, which it will use to develop a masters programme in clinical research. A total of 35 five-year K30 awards were made by NIH during the programme's first year, each funded to the tune of \$200,000 per year.

It is essential to engage students in science while they are still in medical school and thinking about careers. Judith Vaitukaitis, director of NIH's National Center for Research Resources, says the agency plans soon to put out a 'request for applications' for its new medical student training programme. This is modelled on a programme at the University of Pennsylvania, Johns Hopkins University and Washington University. Students will take a year off from medical school and work with a clinical investigator.

Two career development initiatives were also implemented by NIH this year for entry-level and mid-career clinical investigators in patient-oriented research. Like the K30 training awards, these awards stemmed largely from recommendations in the Institute of Medicine and NIH panel reports. For those embarking on a career in clinical research, the K23 award provides between three and five years of support for mentor-supervised research projects. Applicants must have a clinical degree, have completed specialized training and fellowships, and hold an academic appointment at an institution with a general clinical research centre.

The K24 mid-career investigator award provides partial salary support to more experienced physician-scientists. It affords them the protected time they need, away from having to provide patient care, to conduct research and mentor young investigators. It offers between three and five years of support, and individuals are expected to spend up to 50 per cent of their time on award-related work. Vaitukaitis says the NIH funded 85 of 192 K23 applications this year and 81 of 181 K24 applications.

The other part of the training equation is good mentoring. But, with the pressures faced by clinicians today, it is easy to see why mentoring doesn't happen. Poland says he takes mentoring seriously and works a 75–80-hour week to meet his administrative, research and mentoring responsibilities. "I do it because it was done for me," he says.

Unfinished business

Although these NIH initiatives are a step in the right direction, they are fairly small scale. Private foundations (see web table, page 214), as well as pharmaceutical companies, have been long-standing supporters of clinical research. The Howard Hughes Medical Institute, for example, provides support to help launch medical students

In some medical schools, a career in research is not actively promoted or held in such high regard.

and physicians on a research career path. Although the number of MDs and MD/PhDs applying for its postdoctoral research fellowships dropped 45 per cent between 1996 and 1998, it is not known to what extent the NIH initiatives may be having an effect on application rates.

Nathan says that the issue of medical-school debt has barely been touched upon. Physicians emerge from clinical training at 32 or 33 years of age with large debts, and cannot find affordable housing for their families on earnings of \$40,000–60,000 per year without going further into debt, he says. "If I were NIH director I would get a panel together to look at that question." And, given the lack of minority groups in clinical research, which often translates into a lack of minority patients in clinical trials, perhaps any effort should be directed there first, he says.

The \$60-million Clinical Research Enhancement Act, which was reintroduced in the House of Representatives last May, addresses the issue of debt, among other things. "My view is that this is only in part an NIH problem," says Nathan, who feels that changes must be made on many fronts. Nathan would like to see the pulling together of key individuals from the top 10–15 research-oriented medical schools to determine what the deans think about this issue.

Given the apparent disincentives, one might be forgiven for asking why anyone would consider a career in clinical research. Rosenberg has an answer: "I've had the best career I could possibly imagine having had. Sure, it's been hard, and I've lived in two different worlds on more than one occasion, but that's what's been fun about it."

Diane Gershon

Diane Gershon is assistant editor, new technology, at Nature Medicine.

e-mail: techmed@earthlink.net

Further reading

- Wyngaarden, J. B. *N. Engl. J. Med.* **301**, 1254–1259 (1979).
- *Careers in Clinical Research: Obstacles and Opportunities* (Institute of Medicine, National Academy Press, Washington DC, 1994).
- Goldstein, J. L. & Brown, M. S. *J. Clin. Invest.* **99**, 2803–2812 (1997).
- Nathan, D. *J. Am. Med. Soc.* **280**, 1427–1431 (1998).
- Rosenberg, L. E. *Science* **283**, 331–332 (1999).



Nathan: 'knowledge requirement is huge.'